Structural and Functional Studies on Glycosaminoglycan-degrading Enzymes from Bacteria

Yusuke NAKAMICHI

2014

CONTENTS

INTRODUCTION

Glycosaminoglycan (GAG) is a heteropolysaccharide comprising uronic acid and amino sugar residues (1). Based on the sugar composition and mode of glycosidic bonds (*i.e.* 1,3- and 1,4-types), GAGs are categorized as hyaluronan, chondroitin sulfate, heparan sulfate, and heparin (Fig. 1a). Hyaluronan has repeating disaccharide units comprising β-D-glucuronic acid (GlcUA) and *N*-acetyl-D-glucosamine (GlcNAc), while chondroitin sulfate has repeating disaccharide units comprising GlcUA and *N*-acetyl-D-galactosamine (GalNAc). Both disaccharide units of hyaluronan and chondroitin sulfate are linked via 1,3-glycosidic bonds. In contrast, heparan sulfate and heparin have only 1,4-type glycosidic bonds, which comprise repeating disaccharide units with uronic acid residues [α-L-iduronic acid (IdoUA) or β-GlcUA] and amino sugar residues [D-glucosamine (GlcN) or GlcNAc]. Chondroitin sulfate, heparan sulfate, and heparin are frequently sulfated; the sulfation level per repeating disaccharide unit was found to be 1.0 for chondroitin sulfate, 0.8–1.4 for heparan sulfate, and 2.4 for heparin (2). These GAGs play multiple roles in the architecture of animal extracellular matrices and cell growth and differentiation (3). With the exception of hyaluronan, GAGs are present as core protein-bound proteoglycans.

Adhesion of microbes to host cells may be a primary mechanism for residence of normal flora and pathogenic infections. GAGs are typical microbial targets for interactions with host cells, *e.g.* OppA of *Lactobacillus salivarius*, OmcB of *Chlamydia pneumoniae*, M protein of *Streptococcus pyogenes,* and C protein of *Streptococcus agalactiae* bind GAGs of host extracellular matrices (4-7). GAGs are also depolymerized by a variety of bacterial polysaccharide lyases. Based on primary structures, these enzymes are categorized into several families (PL-1–22) in the Carbohydrate-Active enZymes (CAZy) database (8-12). GAG-depolymerizing enzymes include hyaluronate and chondroitin ABC lyases (family PL-8), and heparinase I (PL-13), -II (PL-21), and -III (PL-12), which have differing sequence and substrate specificities (11,12). Polysaccharide lyases depolymerize GAGs through β-elimination reactions and generally produce unsaturated disaccharides containing unsaturated uronic acid with double bonds between C-4 and C-5 positions at nonreducing termini (Fig. 1b) (13). Hyaluronate lyases produced by pathogenic streptococci such as *S. agalactiae*, *Streptococcus pneumoniae*, and *S. pyogenes* are categorized into PL-8 and act as a pathogenic spreading factor; these enzymes are capable of depolymerizing both sulfated and unsulfated chondroitin, as well as hyaluronan (14-18). Researchers have extensively studied the structure and function of streptococcal hyaluronate lyases in order to ultimately facilitate the

development of therapeutic agents for inhibition of the lyases (9,19). On the other hands, degradation and metabolism of the resultant unsaturated disaccharides in streptococci remain to be clarified.

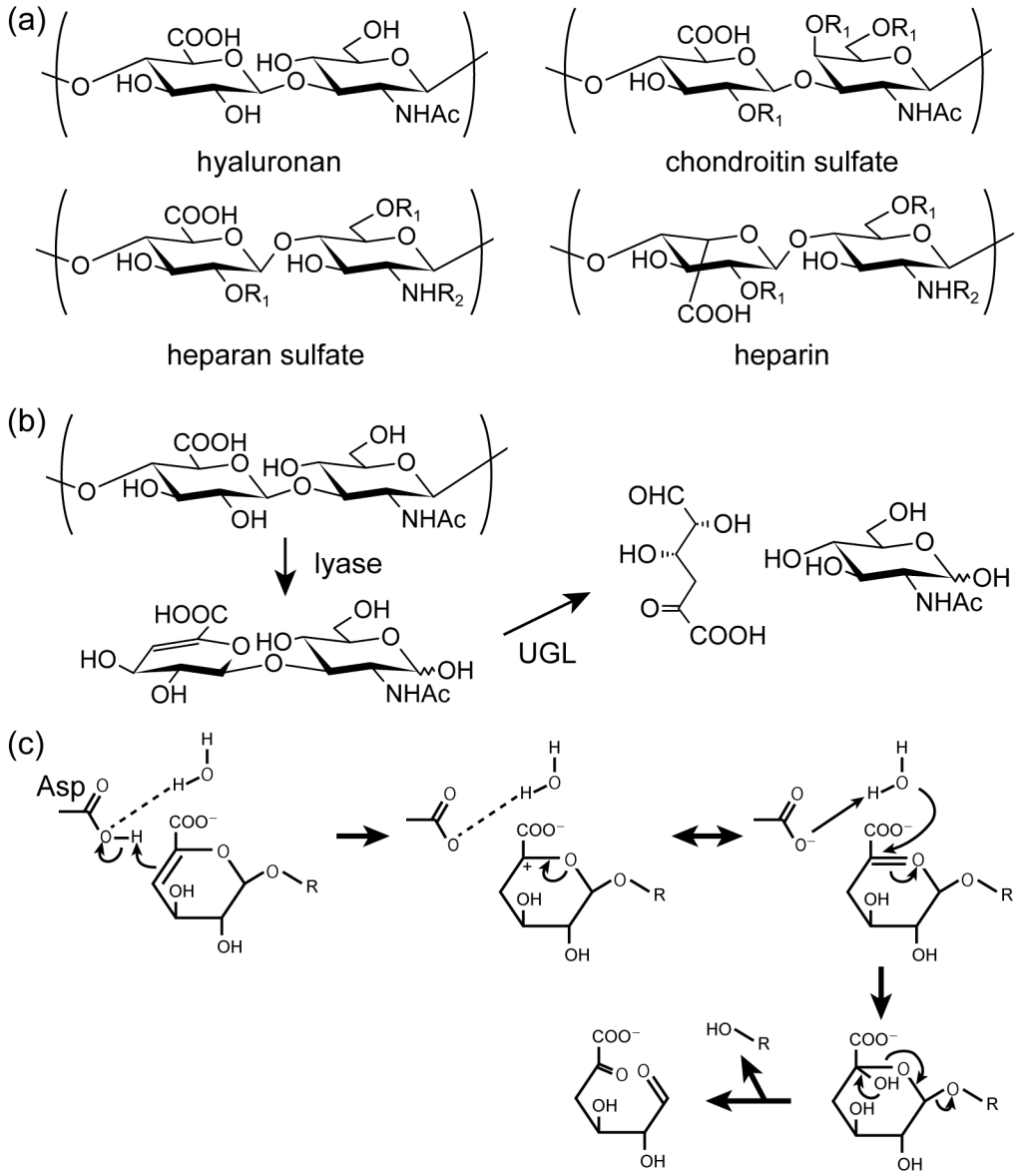

Figure 1. GAGs. (a) Structures of repeating disaccharide units of GAGs. R_1, hydrogen atom or sulfate group; R_2, hydrogen atom, acetyl group, or sulfate group. (b) GAG degradation pathway in *Bacillus* sp. GL1. (c) UGL reaction mechanism.

Unsaturated GAG disaccharides containing unsaturated glucuronic acid (ΔGlcUA) at nonreducing termini, *i.e.* ΔGlcUA-1,3-GalNAc produced by chondroitin lyase and ΔGlcUA-1,3-GlcNAc produced by hyaluronate lyase, are further degraded to monosaccharides by unsaturated glucuronyl hydrolase (UGL), which has been first identified in *Bacillus* sp. GL1 and

classified as a member of the glycoside hydrolase family (GH)-88 in the CAZy database (Fig. 1b) (8,20,21). The structure and function of BacillusUGL have been demonstrated by X-ray crystallography, site-directed mutagenesis, enzyme kinetics, and gas chromatography/mass spectrometry (22,23). Experiments using artificial substrates have also supported its structure–function relationship (24). A novel catalytic mechanism for UGL has previously been proposed as follows: Asp residue in the active site acts as a general acid and base catalyst to protonate the C-4 atom of ΔGlcUA and to deprotonate the water molecule. The deprotonated water molecule attacks the C-5 atom of ΔGlcUA to yield unstable hemiketal; this is followed by spontaneous conversion to an aldehyde (4-deoxy-L-threo-5-hexosulose-uronate) and saccharide through hemiacetal formation and cleavage of the glycosidic bond (Fig. 1c) (23). Although UGLs have been identified and characterized only in soil bacteria such as *Bacillus* sp. GL1 and *Pedobacter heparinus* (formerly *Flavobacterium heparinum*), UGL homologous gene is also present in the genome of the bacteria interacting with host cells such as streptococci, enterococci, lactobacilli, bacteroides, and mycoplasma. Therefore, elucidation of structure–function relationship of bacterial UGLs will contribute to understanding of interaction mode between bacteria and host cells and to development of therapeutic agents for bacterial infection.

In order to clarify the mechanism of interaction between bacteria and host cells such as symbiosis and infection, this thesis focuses on bacterial UGLs and heparan sulfate lyase. CHAPTERs I and II deal with structure and function of bacterial UGL specific for GAG containing 1,3-glycosidic bond (hyaluronan and chondroitin sulfate). Especially, CHAPTER I shows physiological function and crystal structure of streptococcal UGLs, and CHAPTER II reveals structural determinants in the enzymes for recognition of sulfated unsaturated chondroitin disaccharides. CHAPTER III deals with structure–function relationship of bacterial enzymes involved in depolymerization and degradation of GAG containing 1,4-glycosidic bond (heparan sulfate and heparin).

References

1. Ernst, S., Langer, R., Cooney, C. L., and Sasisekharan, R. (1995) Enzymatic degradation of glycosaminoglycans. *Crit. Rev. Biochem. Mol. Biol.* **30**, 387-444
2. Rostand, K. S., and Esko, J. D. (1997) Microbial adherence to and invasion through proteoglycans. *Infect. Immun.* **65**, 1-8
3. Gandhi, N. S., and Mancera, R. L. (2008) The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* **72**, 455-482
4. Martin, R., Martin, C., Escobedo, S., Suarez, J. E., and Quiros, L. M. (2013) Surface glycosaminoglycans mediate adherence between HeLa cells and *Lactobacillus salivarius*

Lv72. *BMC Microbiol.* **13**, 210
5. Stephens, R. S., Koshiyama, K., Lewis, E., and Kubo, A. (2001) Heparin-binding outer membrane protein of chlamydiae. *Mol. Microbiol.* **40**, 691-699
6. Frick, I. M., Schmidtchen, A., and Sjobring, U. (2003) Interactions between M proteins of *Streptococcus pyogenes* and glycosaminoglycans promote bacterial adhesion to host cells. *Eur. J. Biochem.* **270**, 2303-2311
7. Baron, M. J., Bolduc, G. R., Goldberg, M. B., Aupérin, T. C., and Madoff, L. C. (2004) Alpha C protein of group B *Streptococcus* binds host cell surface glycosaminoglycan and enters cells by an actin-dependent mechanism. *J. Biol. Chem.* **279**, 24714-24723
8. Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* **37**, D233-D238
9. Li, S. L., Kelly, S. J., Lamani, E., Ferraroni, M., and Jedrzejas, M. J. (2000) Structural basis of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase. *EMBO J.* **19**, 1228-1240
10. Shim, K. W., and Kim, D. H. (2008) Cloning and expression of chondroitinase AC from *Bacteroides stercoris* HJ-15. *Protein Expr. Purif.* **58**, 222-228
11. Sasisekharan, R., Bulmer, M., Moremen, K. W., Cooney, C. L., and Langer, R. (1993) Cloning and expression of heparinase I gene from *Flavobacterium heparinum*. *Proc. Natl. Acad. Sci. USA* **90**, 3660-3664
12. Su, H. S., Blain, F., Musil, R. A., Zimmermann, J. J. F., Gu, K. F., and Bennett, D. C. (1996) Isolation and expression in *Escherichia coli* of *hepB* and *hepC*, genes coding for the glycosaminoglycan-degrading enzymes heparinase II and heparinase III, respectively, from *Flavobacterium heparinum*. *Appl. Environ. Microbiol.* **62**, 2723-2734
13. Linhardt, R. J., Avci, F. Y., Toida, T., Kim, Y. S., and Cygler, M. (2006) CS lyases: structure, activity, and applications in analysis and the treatment of diseases. *Adv. Pharmacol.* **53**, 187-215
14. Pritchard, D. G., Lin, B., Willingham, T. R., and Baker, J. R. (1994) Characterization of the group B streptococcal hyaluronate lyase. *Arch. Biochem. Biophys.* **315**, 431-437
15. Gase, K., Ozegowski, J., and Malke, H. (1998) The *Streptococcus agalactiae hylB* gene encoding hyaluronate lyase: completion of the sequence and expression analysis. *Biochim. Biophys. Acta Gene Struct. Expr.* **1398**, 86-98
16. Paton, J. C., Andrew, P. W., Boulnois, G. J., and Mitchell, T. J. (1993) Molecular analysis of the pathogenicity of *Streptococcus pneumoniae* - the role of pneumococcal proteins. *Annu. Rev. Microbiol.* **47**, 89-115
17. Berry, A. M., Lock, R. A., Thomas, S. M., Rajan, D. P., Hansman, D., and Paton, J. C. (1994) Cloning and nucleotide sequence of the *Streptococcus pneumoniae* hyaluronidase gene and purification of the enzyme from recombinant *Escherichia coli*. *Infect. Immun.* **62**, 1101-1108
18. Hynes, W. L., Dixon, A. R., Walton, S. L., and Aridgides, L. J. (2000) The extracellular hyaluronidase gene (*hylA*) of *Streptococcus pyogenes*. *FEMS Microbiol. Lett.* **184**, 109-112
19. Mello, L. V., de Groot, B. L., Li, S. L., and Jedrzejas, M. J. (2002) Structure and flexibility of *Streptococcus agalactiae* hyaluronate lyase complex with its substrate - insights into the mechanism of processive degradation of hyaluronan. *J. Biol. Chem.* **277**, 36678-36688
20. Hashimoto, W., Kobayashi, E., Nankai, H., Sato, N., Miya, T., Kawai, S., and Murata, K. (1999) Unsaturated glucuronyl hydrolase of *Bacillus* sp. GL1: novel enzyme prerequisite for metabolism of unsaturated oligosaccharides produced by polysaccharide lyases. *Arch. Biochem. Biophys.* **368**, 367-374

21. Mori, S., Akao, S., Nankai, H., Hashimoto, W., Mikami, B., and Murata, K. (2003) A novel member of glycoside hydrolase family 88: overexpression, purification, and characterization of unsaturated β-glucuronyl hydrolase of *Bacillus* sp. GL1. *Protein Expr. Purif.* **29**, 77-84
22. Itoh, T., Akao, S., Hashimoto, W., Mikami, B., and Murata, K. (2004) Crystal structure of unsaturated glucuronyl hydrolase, responsible for the degradation of glycosaminoglycan, from *Bacillus* sp. GL1 at 1.8 Å resolution. *J. Biol. Chem.* **279**, 31804-31812
23. Itoh, T., Hashimoto, W., Mikami, B., and Murata, K. (2006) Crystal structure of unsaturated glucuronyl hydrolase complexed with substrate. *J. Biol. Chem.* **281**, 29807-29816
24. Jongkees, S. A. K., and Withers, S. G. (2011) Glycoside cleavage by a new mechanism in unsaturated glucuronyl hydrolases. *J. Am. Chem. Soc.* **133**, 19334-19337

CHAPTER I

Gene expression and enzymatic characteristics of streptococcal unsaturated glucuronyl hydrolases

Streptococci such as group B *S. agalactiae*, group nonassigned *S. pneumoniae*, and group A *S. pyogenes* are typical pyogenic and hemolytic pathogens causing severe infections (*e.g.* pneumonia, bacteremia, sinusitis, or meningitis) (1-3). Homologous genes to UGL are found in the genomes of these pathogenic streptococci. These streptococci also produce hyaluronate lyase as a virulence factor, suggesting that UGL involved in degradation of unsaturated GAG disaccharides contributes to their infection in host cells.

This CHAPTER deals with identification and expression analysis of UGL genetic cluster in *S. agalactiae* cells, characterization of streptococcal UGLs, and structural determination of *S. agalactiae* UGL by X-ray crystallography.

Materials and Methods

Homology and Sequence Alignment Analysis. To find novel proteins homologous to *Bacillus* sp. GL1 UGL (BacillusUGL), the *BLAST* and *CLUSTALW* programs were used to search for a sequence similarity and for multiple sequence alignment, respectively. Both programs are available on the DDBJ server (http://www.ddbj.nig.ac.jp).

Microorganisms and Culture Conditions. S. agalactiae strain NEM316 (CIP 82.45) was purchased from the Institut Pasteur. *S. pneumoniae* strain R6 (ATCC BAA-255) and *S. pyogenes* strain M1 GAS SF370 (ATCC 700294) were from the American Type Culture Collection. For genome DNA isolation, the three streptococci were statically grown at 30°C under 5% CO_2 in tryptic soy broth (Difco) containing 5% defibrinated sheep blood (Nippon Bio-Test Laboratories). To investigate gene expression, *S. agalactiae* cells were statically grown at 30°C in tryptic soy broth with or without hyaluronan (Fluka). As a host for plasmid amplification, *Escherichia coli* strain DH5α (Toyobo) cells were cultured at 37°C in LB medium (4) containing sodium ampicillin (0.1 mg ml^{-1}). *E. coli* strain HMS174(DE3)pLysS (Novagen) was used for expressing the three streptococcal UGLs. For expression in *E. coli*, the cells were aerobically precultured at 30°C in LB medium supplemented with sodium ampicillin (0.1 mg ml^{-1}). At ~0.5 turbidity at 600 nm, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the culture at 0.1 mM final concentration, and the cells were further cultured at 16°C for 44 h.

DNA Manipulations. Genomic DNA isolation, subcloning, transformation, and gel electrophoresis were performed as described previously (4). The nucleotide sequences of streptococcal UGL genes were determined by the dideoxy chain termination method using automated DNA sequencer model 377 (Applied Biosystems) (5). Restriction endonucleases and DNA modifying enzymes were from Takara Bio and Toyobo.

Assays for Enzymes and Proteins. Streptococcal UGLs were assayed by monitoring the decrease in absorbance at 235 nm arising from the double bond (molar extinction coefficient ε_{235} = 4,800 M^{-1} cm^{-1}) in the substrate. Standard assay was conducted at 30°C in 20 mM tris (hydroxymethyl) aminomethane (Tris)–HCl (pH 7.5), 0.15 M NaCl, and 0.2 mM substrate. For kinetic assay, the enzyme assay was performed three times using various concentrations (0.1–0.8 mM) of unsaturated chondroitin disaccharide sulfated at C-6 position of GalNAc residue (CΔ6S) (Sigma-Aldrich) as a substrate. One unit of enzyme activity was defined as the amount of enzyme required to degrade 1.0 μmol of substrate per minute by using a cuvette with a 1 cm light path. TLC was performed to investigate the enzyme activity using a high concentration of substrate and enzyme as described previously (6). Protein content was determined by the Bradford method (7), with bovine serum albumin as the standard. For the purified enzyme, protein concentration was estimated by measuring absorbance at 280 nm. Absorbance coefficients of *S. agalactiae* UGL (SagUGL), *S. pneumoniae* UGL (SpnUGL), and *S. pyogenes* UGL (SpyUGL) are 2.30, 2.15, and 2.21 for 1 mg ml^{-1} protein, respectively.

Construction of the Overexpression System. Overexpression systems for SagUGL, SpnUGL, and SpyUGL were constructed in *E. coli* cells as follows. To clone the streptococcal UGL genes, PCR was conducted using each genomic DNA of three streptococci as a template, synthetic oligonucleotide primers, and KOD polymerase (Toyobo). The sequences of oligonucleotides with *Nde*I, *Nco*I, and *Xho*I sites added to their 5' regions for SagUGL, SpnUGL, and SpyUGL are shown in Table 1-1. PCR conditions were as follows: 98°C for 30 sec, 50°C for 2 sec, and 74°C for 2 min with a total of 30 cycles. Final extension was performed by one cycle at 72°C for 10 min. The PCR products were ligated with *Hin*cII-digested pUC119 (Takara Bio) using Ligation High (Toyobo) in accordance with manufacture's protocol, and the resultant plasmids were digested with *Nde*I or *Nco*I and *Xho*I to isolate streptococcal UGL genes. DNA fragments for streptococcal UGL genes were ligated with *Nde*I or *Nco*I and *Xho*I-digested pET21b or pET21d (Novagen). The resultant plasmids were designated pET21b-SagUGL, pET21b-SpnUGL, and pET21d-SpyUGL.

Table 1-1. Primers

Primers for molecular cloning		
SagUGL	forward	5′-GGCATATGATGAAAATAAAACCGGTCAAGG-3′
	reverse	5′-CCCTCGAGTTACCAATAAAGTTCCCAGTCT-3′
SpnUGL	forward	5′-GGCATATGATAAAAAAGGTTACGATTGAAA-3′
	reverse	5′-GGCTCGAGCTACCAATATAGGTTCCAGTCT-3′
SpyUGL	forward	5′-GGCCATGGCACGACCTTTAAAAACCATTGC-3′
	reverse	5′-GGCTCGAGTTACCAGTAAGGGTTCCAGTCT-3′
Primers for site-directed mutagenesis		
D115N	sense	5′-GCATTAGATCACCACAACTTAGGATTTCTTTACACACC-3′
	antisense	5′-GGTGTGTAAAGAAATCCTAAGTTGTGGTGATCTAATGC-3′
D175N	sense	5′-CTATCGCTTAATTATCAATTGCTTACTTAATATCC-3′
	antisense	5′-GGATATTAAGTAAGCAATTGATAATTAAGCGATAG-3′
R236H	sense	5′-CCGTTAAAAGGTGTCACACATCAGGGTTATAGTGATG-3′
	antisense	5′-CATCACTATAACCCTGATGTGTGACACCTTTTAACGG-3′
R236A	sense	5′-CCGTTAAAAGGTGTCACAGCACAGGGTTATAGTGATG-3′
	antisense	5′-CATCACTATAACCCTGTGCTGTGACACCTTTTAACGG-3′

Underlines in primers for molecular cloning and site-directed mutagenesis indicate the restriction and mutation sites, respectively.

Purification. Unless otherwise specified, all operations were carried out at 0–4°C. *E. coli* cells harboring pET21b-SagUGL, pET21b-SpnUGL, or pET21d-SpyUGL were grown in 3.0 liters of LB medium, collected by centrifugation at 6,000 × *g* and 4°C for 5 min, washed with 20 mM potassium phosphate buffer (KPB, pH 7.0), and then resuspended in the same buffer. The cells were ultrasonically disrupted (Insonator model 201M, Kubota) at 0°C and 9 kHz for 20 min, and the clear solution obtained by centrifugation at 20,000 × *g* and 4°C for 20 min was used as the cell extract. SagUGL was purified from the cell extract by anion exchange chromatography (TOYOPEARL DEAE-650M, Tosoh), followed by anion exchange chromatography (MonoQ, GE Healthcare), and finally by gel filtration chromatography (HiLoad 16/60 Superdex 75 pg, GE Healthcare). SpnUGL was purified from the cell extract by anion exchange chromatography (TOYOPEARL DEAE-650M), followed by ammonium sulfate precipitation (50–70% saturation), and finally by gel filtration chromatography (HiLoad 16/60 Superdex 75 pg). SpyUGL was purified from the cell extract by ammonium sulfate precipitation (50–70% saturation) and by gel filtration chromatography (Sephacryl S-200HR, GE Healthcare). The active fractions were combined and dialyzed against 20 mM Tris–HCl (pH 7.5). The dialysate was concentrated to ~8 mg ml^{-1} by ultrafiltration using a Centriprep (molecular mass cut-off, 10 kDa) (Millipore). The concentrate was used as the purified enzyme source.

Determination of Molecular Mass. To determine the molecular weight of streptococcal UGLs, SDS-PAGE (8) and gel filtration chromatography (HiLoad 16/60 Superdex 75 pg) were performed.

Optimal pH and Temperature and Thermal Stability. The experiments were conducted at 30°C by using CΔ6S as the substrate and purified streptococcal UGL. To determine the optimal pH, reactions were performed at 30°C in the following 20 mM buffers: sodium acetate, KPB, and Tris–HCl. To determine the optimal temperature, reactions were performed at various temperatures in 20 mM Tris–HCl (pH 7.5). To determine the thermal stability, after preincubation of the enzyme at various temperatures for 5 min, residual activity was measured at 30°C in 20 mM Tris–HCl (pH 7.5).

Substrate Specificity. Substrate specificity of streptococcal UGLs was determined using various unsaturated saccharides (Fig. 1-1) at 0.2 mM, including CΔ6S, unsaturated gellan tetrasaccharide (ΔGellan) (9), unsaturated hyaluronan disaccharide (ΔHA), sulfate group-free unsaturated chondroitin disaccharide (CΔ0S), unsaturated chondroitin disaccharide sulfated at the C-4 position of GalNAc residue (CΔ4S), unsaturated chondroitin disaccharide sulfated at the C-2 position of ΔGlcUA residue (CΔ2'S), unsaturated chondroitin disaccharide sulfated at the C-2 position of ΔGlcUA residue and the C-4 position of GalNAc residue (CΔ2'S4S), unsaturated chondroitin disaccharide sulfated at the C-2 position of ΔGlcUA residue and the C-6 position of GalNAc residue (CΔ2'S6S), unsaturated chondroitin disaccharide sulfated at the C-4 and C-6 positions of GalNAc residue (CΔ4S6S), and unsaturated chondroitin disaccharide sulfated at the C-2 position of ΔGlcUA residue and the C-4 and C-6 positions of GalNAc residue (CΔ2'S4S6S). These unsaturated disaccharides were from Seikagaku Kogyo.

DNA Microarray. Total RNAs for DNA microarray were prepared as follows. *S. agalactiae* cells grown at exponential growth phase (turbidity at 600 nm, 0.7) were collected at 6,000 × *g* and 4°C for 5 min and treated with RNAprotect bacteria reagent (Qiagen) for stabilization of RNA. Based on the manufacturer's instructions, total RNAs were extracted with an RNeasy kit (Qiagen) and treated with DNase (RNase-free DNase set, Qiagen) supplied with the kit. The genome sequence and open reading frame prediction for *S. agalactiae* were provided by the Genome Sequence Project (1) (gib.genes.nig.ac.jp/single/index.php?spid=Saga_NEM316). DNA microarray from NimbleGen Systems includes 2,094 target genes fixed on the glass slide by fixing at least 20 unique probes consisting of 60-mer synthetic oligonucleotides for each gene. Biotin labeling, fragmentation, and hybridization were done by NimbleGen Systems. The arrays were scanned with an Axon Genepix 4000B scanner at 532 nm and 5-μm resolution and analyzed by quantile normalization and robust multiarray averaging (10). These normalized data were processed with NANDEMO analysis 1.0.0 software (Roche Applied Science). Student's *t* test for analyzing the

mean log ratios of two samples and subsequent Bonferroni adjustment for multiple testing (2,094 open reading frames on arrays) were applied as a rigorous criterion for significantly changed signal intensity. Changes with $p < 0.05$ were considered statistically significant.

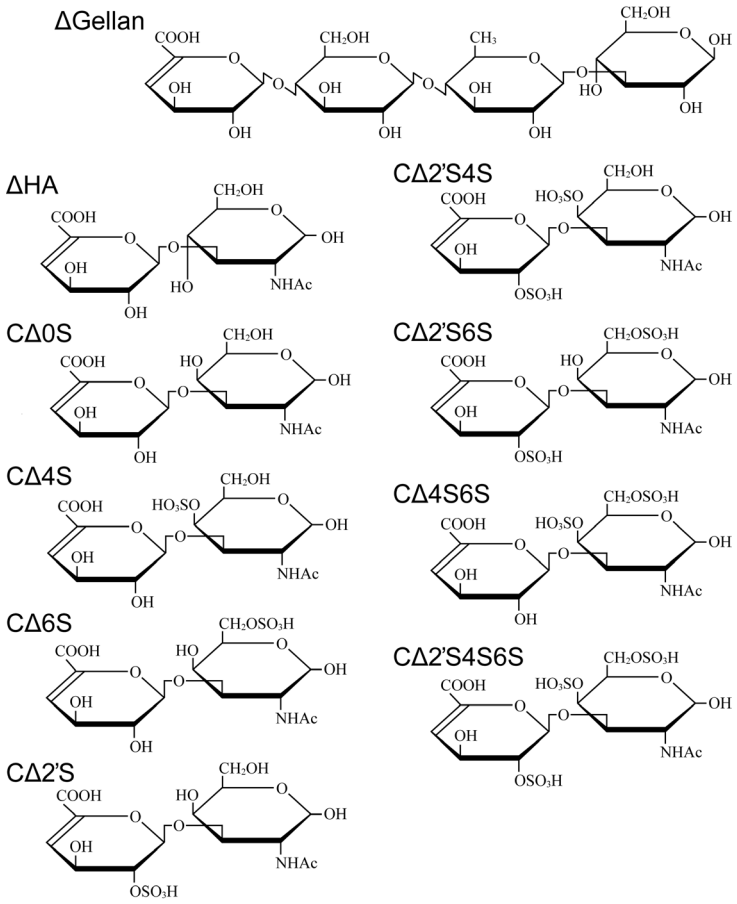

Figure 1-1. Structures of unsaturated oligosaccharides.

Western Blotting. Antibodies to the purified SpyUGL were raised in a rabbit, and the serum was used as polyclonal antibodies against streptococcal UGLs. To investigate SagUGL expression in *S. agalactiae*, the cell extracts were subjected to SDS-PAGE, followed by Western blotting using anti-SpyUGL antibodies as described previously (11). Anti-IgG conjugated with horseradish peroxidase (GE Healthcare) and a POD immunostaining kit (Wako Pure Chemical Industries) were used to visualize the protein band.

Crystallization and Structure Determination. To determine the three-dimensional structure of SagUGL, the purified enzyme was crystallized by the hanging drop vapor diffusion method. The 3 μl of proteins (8 mg ml^{-1} in 20 mM Tris–HCl, pH 7.5) was mixed with equal volume of a reservoir solution (100 mM MES–NaOH, pH 6.0, 1.9 M ammonium sulfate). The protein solution was then

incubated at 20°C, and single crystals grew up for ~1 month. The crystal was flash frozen under a cold nitrogen gas stream at –173°C after soaking in mother liquor containing 20% glycerol as a cryoprotectant. X-ray diffraction data were collected at λ = 0.7000 Å using a Jupiter 210 CCD detector (Rigaku) at BL38B1 station of SPring-8 (Hyogo, Japan). The data were processed and scaled up to 1.75 Å with the *HKL2000* program (12). The structure was determined by molecular replacement with the *Molrep* program (13) supplied in a *CCP4* program package (14) by using coordinates of BacillusUGL [Protein Data Bank (PDB) ID: 1VD5] as an initial model. Structure refinement was conducted with the *Refmac5* program (15). Randomly selected 5% reflections were excluded from refinement and used to calculate R_{free}. After each refinement cycle, the model was adjusted manually using the *winCoot* program (16). Final model quality was checked with the *PROCHECK* program (17). Model structure of SagUGL binding to CΔ4S was built using coordinates of chondroitin-4 sulfate (PDB ID: 1C4S), followed by energy minimization with the *Refmac5* program (15). Figures for the protein structure were prepared using the *PyMol* program (18). Coordinates used in this CHAPTER were taken from the Research Collaboratory for Structural Bioinformatics (RCSB) PDB (19). The atomic coordinates and structure factors of SagUGL (ID: 2ZZR) have been deposited in the RCSB PDB (http://www.rcsb.org/).

Site-directed Mutagenesis. To substitute Asp-115, Asp-175, Arg-236, and Arg-236 of SagUGL with Asn, Asn, His, and Ala, respectively, eight oligonucleotides were synthesized as described in Table 1-1. Site-directed mutagenesis was performed using the plasmid pET21b-SagUGL as a template and synthetic oligonucleotides as sense and antisense primers by employing methods of a QuikChange site-directed mutagenesis kit (Stratagene), except that KOD-Plus polymerase (Toyobo) was used for PCRs. The resultant plasmids with a mutation were designated pD115N, pD175N, pR236H, and pR236A, respectively. The mutations were confirmed by DNA sequencing. *E. coli* host strain [HMS174(DE3)] cells were transformed with the each plasmid.

Results and Discussion

Streptococcal Proteins Similar to GH-88 UGL

To find streptococcal proteins homologous to BacillusUGL (GenBank™ accession number AB019619), the *BLAST* program was used to search for sequence similarity against the streptococcal genome databases. As described previously (20), *S. pneumoniae* and *S. pyogenes* have an open reading frame similar to BacillusUGL: *S. pneumoniae* spr0292 (Uniprot ID, Q8DR77; 396 residues, 46.0 kDa) and *S. pyogenes* Spy0632 (Uniprot ID, Q9A0T3; 399 residues, 46.1 kDa). In *S.*

pneumoniae, the gene for UGL homolog (spr0292) is located near the hyaluronate lyase (spr0286) gene (*hysA*) in the genome (2) (Fig. 1-2a), suggesting that the two proteins are cooperatively involved in degrading GAGs. A UGL homologous protein is also encoded in the genome of *S. agalactiae*: *S. agalactiae* Gbs1889 (Uniprot ID, Q8E372; 398 residues, 46.6 kDa). Gbs1889 showed 40.3% identity with BacillusUGL in a 352-amino acid overlap, and the three streptococcal proteins are mutually similar (identity, over 69.1%) (Fig. 1-2d). Based on high identity, the three proteins are assigned to unsaturated glucuronyl hydrolase on each genome database. In the three pathogenic streptococci, hyaluronate lyases are believed to function as a virulent factor (21-23), suggesting that UGL and hyaluronate lyase play an important role in the bacterial infection process. Because SpnUGL was predicted to be localized in the cytoplasm by the *PSORT* program (www.psort.org/) and BacillusUGL is demonstrated to be expressed in the intracellular fraction (6), streptococcal UGLs are probably cytoplasmic enzymes. Interestingly, the UGL genes of the three streptococci are located in a gene cluster for the putative phosphotransferase system to incorporate amino sugar (GalNAc), a component of GAGs (Fig. 1-2a, b). The phosphotransferase system consists of several domains such as enzymes IIA (E-IIA), IIB (E-IIB), IIC (E-IIC), and IID (E-IID) and incorporate solutes across the cytoplasmic membrane through phosphate addition to the substrates (24). The streptococcal genetic organization around the UGL gene strongly suggests that UGL could universally function in pathogenic bacteria as a virulent factor responsible for the complete degradation of GAGs after depolymerization by polysaccharide lyase, followed by import of the resultant unsaturated disaccharides through phosphotransferase system. Furthermore, in the genome of the three streptococci, the putative genes coding for 5-keto-D-gluconate reductase, 2-keto-3-deoxygluconate kinase, and 2-keto-3-deoxy-6-phosphogluconate aldolase involved in the metabolism of 2-keto-3-deoxygluconate are located upstream of the UGL gene (Fig. 1-2a, b). 5-Keto-D-gluconate reductase catalyzes a reversible reduction of 5-keto-D-gluconate to form D-gluconate (25). When 5-keto-D-gluconate reductase can also catalyze the reduction reaction of 4-deoxy-L-threo-5-hexosulose-uronate, a reaction product by UGL, the resultant product probably becomes a substrate for 2-keto-3-deoxygluconate kinase and is assimilated in the 2-keto-3-deoxygluconate metabolism pathway. The experimental data are required for elucidating the predicted pathway.

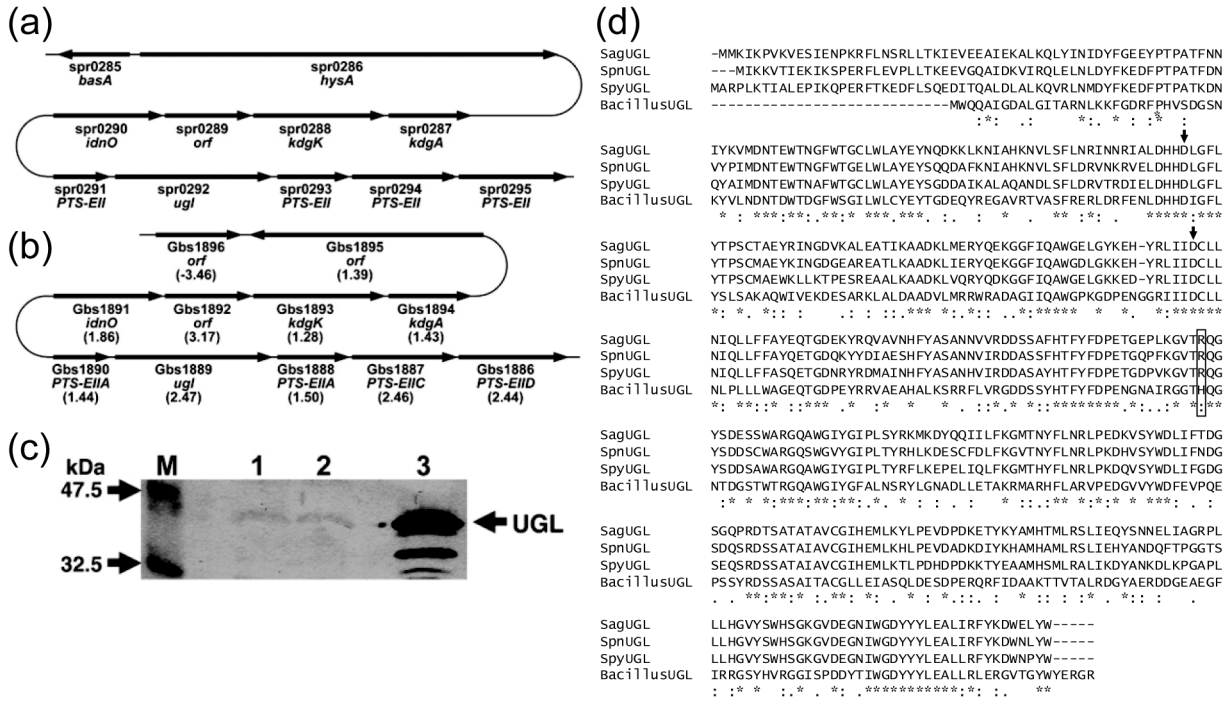

Figure 1-2. UGL gene and protein in streptococci. (a) *S. pneumoniae* gene cluster. (b) *S. agalactiae* gene cluster. Except for Gbs1896, the values in parentheses indicate the increased level of each gene expression in *S. agalactiae* cells grown in the presence of hyaluronan compared with that in the absence of the polysaccharide based on the DNA microarray. The gene expression level except for Gbs1895 between the two bacterial cells was significantly different ($p < 0.05$). Details regarding the gene assignment have been described in the text. The open reading frame indicates a hypothetical protein. (c) Expression of the UGL protein in *S. agalactiae* cells. The cells grown in the presence and absence of hyaluronan were subjected to Western blotting using anti-SpyUGL antibodies. *Lane M*, prestained protein markers with molecular masses of 47.5 and 32.5 kDa; *lane 1*, *S. agalactiae* cells grown in the absence of hyaluronan; *lane 2*, *S. agalactiae* cells grown in the presence of hyaluronan; *lane 3*, purified SagUGL. Purified SagUGL in *lane 3* included faint degraded products. (d) Amino acid sequence alignment of streptococcal and bacillus UGLs. The *arrows* indicate the catalytically important Asp residues. Possible streptococcal UGL arginine residues (SagUGL, Arg-236) involved in binding to sulfate group of CΔ4S are *boxed*. Identical and similar amino acid residues among the four proteins are denoted by *asterisks* and *dots*, respectively.

Enzyme Characteristics

To examine whether streptococcal UGL-homologous proteins show unsaturated glucuronyl hydrolase activity, the proteins expressed in *E. coli* were purified (Table 1-2 and Fig. 1-3a) and subjected to the enzyme assay. In the presence of each protein, the absorbance at 235 nm derived from the substrate in the reaction mixture decreased, indicating that the three proteins exhibit UGL enzyme activity. Gbs1889, spr0292, and Spy0632 are designated SagUGL, SpnUGL, and SpyUGL, respectively. The properties of the purified enzymes are as follows.

Table 1-2. Purification of streptococcal UGLs from *E. coli* cells

Step[a]	Protein (mg)	Total activity (kU)	Specific Activity (U mg^{-1})	Yield (%)	Fold
SagUGL					
Cell extract	1,900	15	7.9	100	1.0
TOYOPEARL DEAE-650M	380	6.7	18	45	2.3
MonoQ	84	8.5	100	57	13
Superdex 75 pg	61	6.8	110	45	14
SpnUGL					
Cell extract	1,500	15	10	100	1.0
TOYOPEARL DEAE-650M	370	8.8	24	59	2.4
Ammonium sulfate (50–70% sat.)	75	2.0	27	13	2.7
Superdex 75 pg	40	1.4	35	9.3	3.5
SpyUGL					
Cell extract	2,000	11	5.5	100	1.0
Ammonium sulfate (50–70% sat.)	690	9.5	14	86	2.5
Sephacryl S-200 HR	150	2.0	13	18	2.4

[a] Purification steps are detailed in the section *Materials and Methods*.

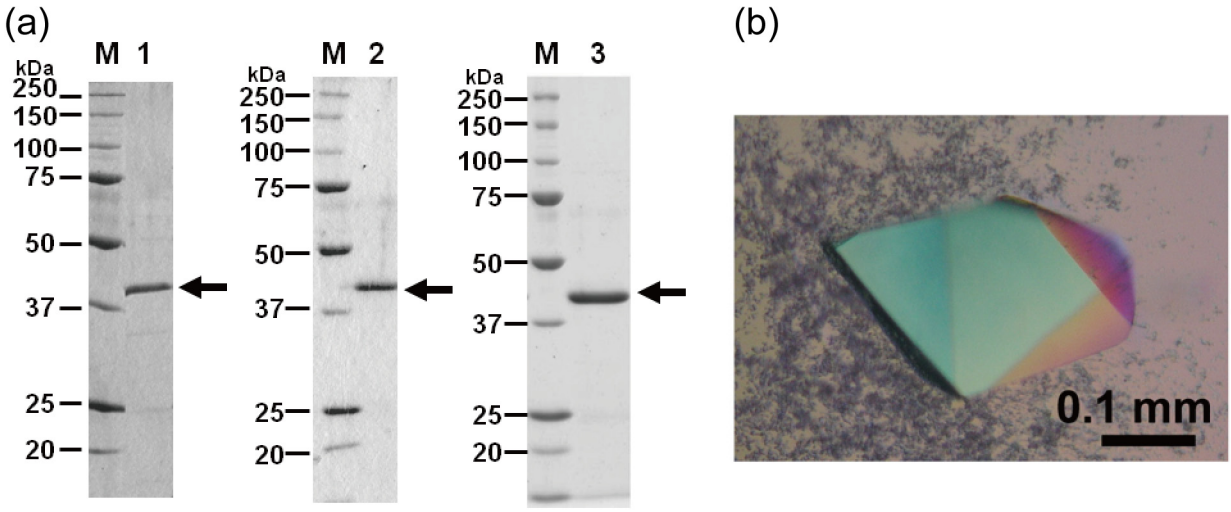

Figure 1-3. Purification and crystallization of streptococcal UGLs. (a) SDS-PAGE profile. *Lane M,* molecular weight standards (from top): synthetic polypeptides with molecular weights of 250,000, 150,000, 100,000, 75,000, 50,000, 37,000, 25,000, and 20,000; *lane 1*, SagUGL; *lane 2*, SpnUGL; *lane 3*, SpyUGL. (b) SagUGL crystal.

Molecular Weight. The molecular masses of SagUGL, SpnUGL, and SpyUGL determined by SDS-PAGE were 43 kDa (Fig. 1-3a). These values were almost comparable with the theoretical one (46 kDa) deduced from the predicted amino acid sequences of the enzymes. In gel filtration chromatography on HiLoad 16/60 Superdex 75 pg, all of the enzymes were eluted between bovine

serum albumin (67 kDa) and ovalbumin (43 kDa), indicating that the enzymes are monomeric.

pH and Temperature. SagUGL, SpnUGL, and SpyUGL were most active at approximately pH 5.5 in 20 mM KPB and at 37°C (Fig. 1-4a, b). About 50% of enzyme activity was observed above 45°C for 5 min in 20 mM Tris–HCl (pH 7.5) (Fig. 1-4c).

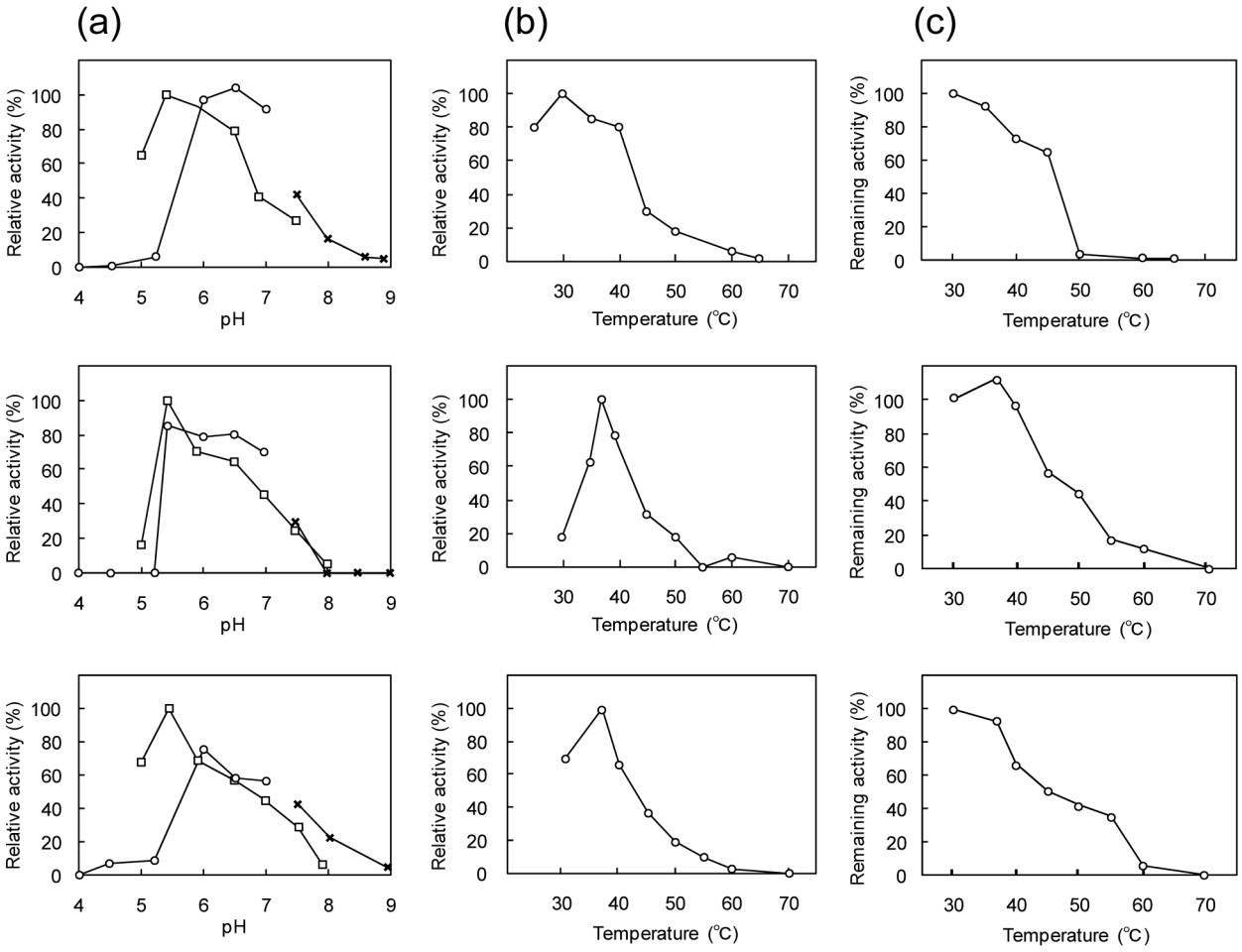

Figure 1-4. Optimal pH and temperature and thermal stability. *Upper,* SagUGL; *center*, SpnUGL; *lower*, SpyUGL. (a) Optimal pH. Experiments were performed at 30°C using the following buffers: 20 mM sodium acetate (○), 20 mM KPB (□), and 20 mM Tris-HCl (×). Maximum activity was relatively taken as 100%. (b) Optimal temperature. Reactions were carried out at various temperatures in 20 mM Tris-HCl (pH 7.5). Maximum activity was relatively taken as 100%. (c) Thermal stability. After preincubation for 5 min at various temperatures, residual activity was measured at 30°C in 20 mM Tris-HCl (pH 7.5). Activity of the enzyme preincubated at 30°C was taken as 100%.

Effect of Reducing Agents, Monosaccharides, and Metals. Table 1-3 shows the effects of different chemicals on enzyme. Streptococcal UGLs were significantly inhibited by uronic acid such as GlcUA or galacturonic acid, suggesting that these saccharides were accommodated in the

active site of the enzymes. Metal ions and thiol reagents at 1 mM had no significant effects on enzyme activity.

Table 1-3. Effects of reducing agents, monosaccharides, and metals on streptococcal UGL activity

Chemical	Concentration (mM)	Relative activity (%)		
		SagUGL	SpnUGL	SpyUGL
None		100	100	100
Reducing agents				
dithiothreitol	1	115	98	104
reduced glutathione	1	138	125	94
2-mercaptoethanol	1	118	90	80
Sugars				
D-GlcUA	5	1.6	3.3	2.2
D-glucose	5	120	100	60
D-mannose	5	100	100	78
D-galactose	5	120	110	79
L-rhamnose	5	110	110	65
D-glucosamine	5	100	130	98
D-sucrose	5	110	100	81
D-xylose	5	110	110	67
D-galacturonic acid	5	0.37	1.1	ND[a]
2-deoxyglucose	5	97	100	89
L-fucose	5	95	100	50
GalNAc	5	90	97	95
GlcNAc	5	83	98	96
Metals				
$CoCl_2$	1	45	ND	ND
$CaCl_2$	1	92	101	97
$MnCl_2$	1	73	95	23
$MgCl_2$	1	91	101	110
$ZnCl_2$	1	ND	ND	6.1
$AlCl_3$	1	35	140	ND
LiCl	1	76	94	130
NaCl	1	78	96	130
KCl	1	78	97	130

[a] ND, not detected; 0.2% < of the control (none).

Substrate Specificity. Table 1-4 shows substrate specificity of SagUGL, SpnUGL, and SpyUGL. CΔ6S was degraded most efficiently by all the streptococcal UGLs. Unlike BacillusUGL (20), the unsaturated gellan tetrasaccharide produced from bacterial biofilm was almost inert on the

streptococcal UGLs. Although the degradation level was extremely low, SagUGL and SpyUGL can act on unsaturated chondroitin disaccharide with sulfate group(s). To confirm no activity of BacillusUGL and SpnUGL on CΔ4S, TLC was performed using a large amount of the enzymes. CΔ4S was degraded by SpnUGL but not by BacillusUGL, whereas both enzymes acted on CΔ6S (Fig. 1-5a–c). Thus, there is a significant difference in the substrate specificity between streptococcal and bacillus UGLs. Streptococcal UGLs seem to have a tendency to act on sulfated substrates.

Table 1-4. Substrate specificity of streptococcal and bacillus UGLs

Substrate	Relative activity (%)			
	SagUGL	SpnUGL	SpyUGL	BacillusUGL[a]
CΔ6S	100	100	100	21
CΔ0S	3.4	16	10	100
ΔHA	0.61	2.0	1.1	62
CΔ2'S	0.011	0.31	0.34	17
ΔGellan	0.050	ND[b]	0.11	78
CΔ4S	0.0090	ND	0.076	ND
CΔ2'S6S	0.0044	ND	0.051	ND
CΔ4S6S	0.019	ND	0.014	ND
CΔ2'S4S	0.0075	ND	ND	ND
CΔ2'S4S6S	ND	ND	ND	ND

[a] Cited from Ref. (20)
[b] ND, not detected; 0.002% < toward CΔ6S.

Kinetics. The K_m values of SagUGL, SpnUGL, and SpyUGL for CΔ6S were 0.54 ± 0.19, 0.18 ± 0.04, and 0.39 ± 0.10 mM, respectively, suggesting that streptococcal UGLs mutually show similar affinity with the substrate. On the other hand, the k_{cat} value of SagUGL is the highest (24 ± 4 s^{-1}) in comparison with those of SpnUGL (1.3 ± 0.1 s^{-1}) and SpyUGL (4.0 ± 0.7 s^{-1}), indicating that SagUGL is relatively active on the substrate rather than the other two enzymes.

To date, two UGLs from *Bacillus* sp. GL1 (6) and *P. heparinus* (26) have been characterized, although *P. heparinus* UGL is specific to unsaturated disaccharides from heparin/heparan sulfate. The molecular weight, optimal reaction conditions, and thermal stability of streptococcal UGLs were comparable with those of bacillus and *P. heparinus* UGLs. Except for the reaction turnover (k_{cat}), there is no significant difference in enzyme characteristics among streptococcal UGLs. Distinct from BacillusUGL, the streptococcal UGLs were capable of degrading unsaturated disaccharides derived from mammalian sulfated GAGs. This substrate specificity probably enables streptococci to readily invade mammalian cells through the degradation of extracellular sulfated

GAGs. Because SagUGL showed the highest activity, the expression and structural analysis of SagUGL is focused on hereafter.

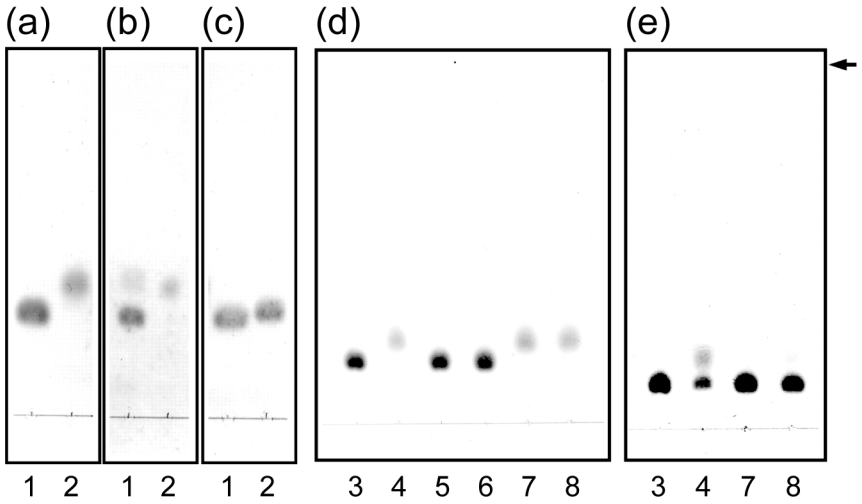

Figure 1-5. Degradation of sulfated unsaturated chondroitin disaccharides by UGLs. (a)−(c) Degradation of CΔ4S by SpnUGL but not by BacillusUGL. The 10 mM substrate CΔ4S or CΔ6S incubated at 30°C with 2 mg ml^{-1} BacillusUGL (a) or SpnUGL (b) or without enzyme (c) was subjected to TLC analysis. *Lane 1*, CΔ4S; *lane 2*, CΔ6S. (d), (e) SagUGL mutant enzyme activity by TLC analysis. *Lane 3*, no enzyme; *lane 4*, wild-type enzyme; *lane 5*, D115N; *lane 6*, D175N; *lane 7*, R236H; *lane 8*, R236A. (d) Degradation of CΔ6S. Reactions were performed at 30°C for 30 min using 10 mM substrate and 0.1 mg ml^{-1} enzyme. (e) Degradation of CΔ4S. Reactions were carried out at 30°C for 60 min using 10 mM substrate and 2 mg ml^{-1} enzyme. The *arrow* indicates the position to which the developing solvent reached.

Expression of UGL in *Streptococcus*

To investigate UGL gene expression in *S. agalactiae* cells, DNA microarray was conducted using the streptococcal cells grown with or without hyaluronan. Hyaluronan induced the expression of the genetic cluster for UGL and the phosphotransferase system (Fig. 1-2b), although the increased level was not very high (less than 3-fold). Hyaluronan also slightly up-regulated the transcriptional level of putative genes involved in 2-keto-3-deoxygluconate metabolism. The SagUGL protein expression in both *S. agalactiae* cells grown with or without hyaluronan was confirmed by Western blotting with anti-SpyUGL antibodies (Fig. 1-2c). These results suggest that SagUGL is constitutively expressed, but the slight increase in its expression level occurs in the presence of hyaluronan.

Structure Determination of SagUGL

To investigate structure determinants for different substrate specificity between streptococcal and bacillus UGLs, X-ray crystallography of SagUGL was performed. SagUGL yielded single

crystals suitable for X-ray analysis (Fig. 1-3b). The crystal structure of SagUGL was determined at 1.75 Å resolution via molecular replacement using BacillusUGL structure (PDB ID: 1VD5) as an initial model. Data collection and model refinement statistics are summarized in Table 1-5. The final model contains one monomer enzyme starting from Met-2 to Trp-398, three sulfate ions, three glycerol molecules, and 407 water molecules. The N-terminal amino acid (Met-1) could not be assigned because of the thin electron density for the corresponding residue in the $2F_o-F_c$ map. Most (89.6%) nonglycine residues lie within the most favored regions of the Ramachandran plot as defined in the *PROCHECK* program (17), and no residues lie within the disallowed regions.

Crystal Structure of SagUGL

A ribbon diagram of SagUGL is shown in Fig. 1-6a. The enzyme is ~45 × 45 × 40 Å in size. SagUGL consists of the α_6/α_6-barrel structure seen in the six-hairpin glycosylases superfamily of the SCOP database (scop.mrc-lmb.cam.ac.uk/scop/), which contains glucoamylases (27), cellulase catalytic domains (28), *N*-acyl-D-glucosamine epimerase (29), and unsaturated galacturonyl hydrolase (30). The root mean square deviation (r.m.s.d.) value for overall 370 C^{α} atoms between SagUGL and BacillusUGL is 1.5 Å, indicating that their structures are essentially the same. According to the designation of the secondary structure element of BacillusUGL (31), SagUGL contains twelve α helices (H1–H12) and seven β strands (SA1-SA2, SB1-SB2, and SC1-SC3). Helices H1, H3, H5, H7, H9, and H11 form the outer barrel, and the other helices, H2, H4, H6, H8, H10, and H12, form the inner barrel. The N-terminal 27 residues are peculiar to SagUGL. This N-terminal region was observed obviously in the crystal structure and lies along with outer barrel (Fig. 1-6a). One short α helix (H0) is additionally formed in the N-terminal loop region of SagUGL.

Active Site Structure

To determine the structure–function relationship, structural comparison between SagUGL and ligand-bound BacillusUGL was conducted to identify the structural determinants for different substrate specificity between streptococcal and bacillus UGLs. In BacillusUGL, the central cleft surrounded by aromatic and positively charged residues functions as the active site (32). Crystal structures of BacillusUGL in complex with CΔ0S, ΔHA, and ΔGellan reveal that UGL is unusual in that no atom forms hydrogen bonds to the substrate glycosidic oxygen to be hydrolyzed (32,33). Based on the complex structures and reaction product analysis using ^{18}O water, the novel catalytic mechanism to hydrolyze the glycosidic bond in which Asp-149 attacks the C4–C5 double bond of ΔGlcUA was proposed (32). Asp-149 acts as a general acid and base catalyst to protonate the ΔGlcUA C-4 atom and to deprotonate the water molecule. The deprotonated water molecule attacks

Table 1-5. Statistics of SagUGL for X-ray diffraction data and structure refinement

Data collection	
Wavelength (Å)	0.7000
Resolution limit (last shell)[a] (Å)	50.00–1.75 (1.81–1.75)[a]
Space group	$P3_221$
Unit cell parameters (Å)	$a = b = 116.1$, $c = 78.1$
Measured reflections	774,026
Unique reflections	61,457
Completeness (%)	100 (100)
I/σ (I)	33.2 (3.89)
R_{merge}	0.056 (0.38)
Refinement	
R_{work}	0.182
R_{free}	0.214
No. of molecules/asymmetric unit	1
No. of non-hydrogen atoms	
Protein	3,329
Sulfate ions	15
Glycerols	18
Water molecules	407
Average isotropic B-factor (Å^2)	
Protein	19.3
Sulfate ions	49.4
Glycerols	49.7
Water molecules	33.7
Root mean square deviation from ideal	
Bond lengths (Å)	0.007
Bond angles (deg)	1.001
Ramachandran plot (%)	
Favored region	98.0
Allowed region	2.0
Outlier region	0

[a] Data on highest shells are in parenthesis.

the ΔGlcUA C-5 atom to yield unstable hemiketal; this is followed by spontaneous conversion to an aldehyde (4-deoxy-L-threo-5-hexosulose-uronate) and saccharide through hemiacetal formation and cleavage of the glycosidic bond. BacillusUGL strictly recognizes ΔGlcUA at subsite −1 rather than the saccharide at subsite +1 (33). It is reasonable in that the enzyme acts on various unsaturated oligosaccharides but not on saturated oligosaccharides.

Most residues in the active site cleft of SagUGL and BacillusUGL are well conserved in both primary and tertiary structures (Figs. 1-2d and 1-6b, c). In Fig. 1-6b, side chain conformation of the catalytic residue (Asp-149 in BacillusUGL and Asp-175 in SagUGL) differs between the two. The side chain of Asp-149 moves upon substrate binding (32), and conformation of SagUGL Asp-175 is

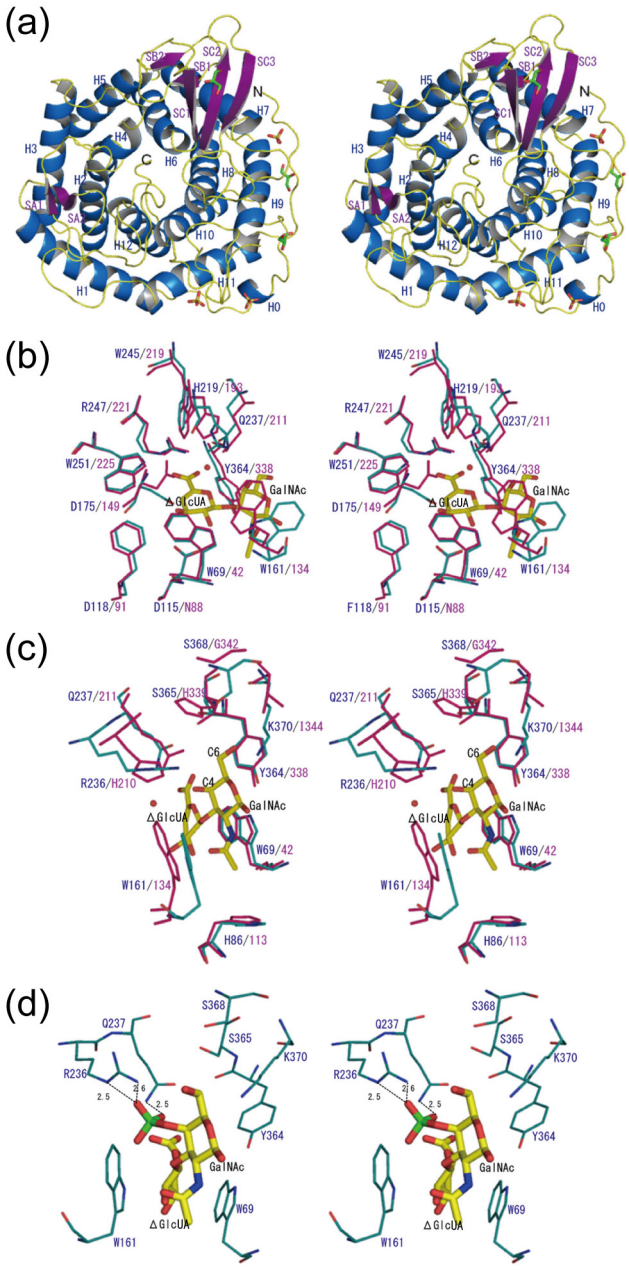

Figure 1-6. Crystal structure of SagUGL (stereo-diagram). (a) The colors *blue*, *purple*, and *yellow* denote the secondary structure elements α-helices, β-sheets, and loops, respectively. Bound sulfate ions and glycerol molecules are represented as stick models. (b) and (c) *Cyan* represents SagUGL, and *pink* denotes BacillusUGL D88N in complex with CΔ0S. Amino acid residues at subsite −1 (b) and +1 (c) are represented in the stick models. Substrate bound in D88N is also shown in the stick models. Carbon, oxygen, and nitrogen atoms of bound saccharides are colored by *yellow*, *pink*, and *blue*, respectively. (d) Energy-minimized model of SagUGL binding to CΔ4S. Arg-236 and Gln-237 bind to the sulfate group of CΔ4S through the formation of hydrogen bonds (*broken lines*). Hydrogen bond lengths are indicated. Carbon, oxygen, nitrogen, and sulfur atoms of bound saccharides are colored by *yellow*, *pink*, *blue*, and *green*, respectively.

similar to that of Asp-149 of substrate-free BacillusUGL. Side chain conformation of SagUGL Trp-161 and BacillusUGL Trp-134 is also different. This side chain stacks with sugar ring of various substrates at subsite +1 in BacillusUGL by altering its conformation. Therefore, the conformational differences in the two amino acid residues seen in Fig. 1-6b may not be effective for the difference in enzyme characteristics of SagUGL and BacillusUGL. On the other hand, amino acid residues at subsite +1 are somewhat different between the two (Fig. 1-6c). His-210 forming a hydrogen bond to *O*-4 of GalNAc residue in BacillusUGL in complex with unsaturated chondroitin disaccharide (32) is substituted with Arg-236 in SagUGL (Figs. 1-2d and 1-6c). His-339 interacting with C-6 of GalNAc residue in BacillusUGL is also replaced by Ser-365 in SagUGL. In addition to these substitutions of residues interacting with the substrate, other differences are observed at subsite +1. BacillusUGL Gly-342 and Ile-344 have no interaction with the substrate, but substitution of these residues with Ser-368 and Lys-370 residues in SagUGL may affect substrate binding, because their side chains are directed to C-6 atom of GalNAc and can interact with the substrate (CΔ6S) when C-6 is sulfated (Fig. 1-6c). Although the overall electron density of SagUGL was good enough to determine the precise position of amino acid residues, that for a loop containing Arg-236 and Gln-237, was slightly poor to determine the exact side chain structure of some residues in the loop. This feature was the same in the structure determined using another crystal belonging to the space group of $P2_1$ (data not shown). These observations suggest conformational disorder of the loop. Because the mutant enzyme of BacillusUGL that substituted Gln-211, a neighbor residue of His-210 in BacillusUGL (Arg-236 of SagUGL), with alanine shows lower k_{cat} and higher K_m (32), the loop disorder is considered to regulate SagUGL enzymatic activity. SagUGL Arg-236, Ser-365, Ser-368, and Lys-370 are common to the other two SpnUGL and SpyUGL (Fig. 1-2d). The loop disorder and/or differences in residues (Arg-236 versus His-210, Ser-365 versus His-339, Ser-368 versus Gly-342, and Lys-370 versus Ile-344) are therefore attributed to the difference seen in substrate specificity between streptococcal and bacillus UGLs, *i.e.* BacillusUGL prefers unsulfated substrate and cannot degrade CΔ4S, although SagUGL showed a high specificity for CΔ6S and acted on CΔ4S.

Amino Acid Residue Binding to Sulfate Group

Asp-149 of BacillusUGL is important for catalysis, and substitution of this residue with Asn reduced k_{cat} by 1,000-fold compared with that of the wild-type enzyme (32). Mutant enzyme replaced Asp-88 by Asn also yields a decreased k_{cat} value (1/10,000) and an increased K_m value (10-fold), indicating the importance of this residue for catalytic reaction. To investigate roles of

corresponding residues (Asp-175 and Asp-115) in SagUGL, mutants SagUGL D175N and D115N were constructed by site-directed mutagenesis. Compared with the wild-type enzyme with k_{cat} of 24 ± 4 s^{-1} and K_m of 0.54 ± 0.19 mM, substitution of Asp-175 with Asn in SagUGL decreased k_{cat} 0.038 ± 0.067 s^{-1} but had no effect on K_m (0.53 ± 0.18 mM). On the other hand, it was very difficult to determine k_{cat} and exactly by kinetic study of SagUGL D115N because K_m of SagUGL D115N is much higher than 1 mM. Enzyme assay with higher concentration of unsaturated substrate is difficult because absorbance at 235 nm derived from the C4–C5 double bond of the substrate is very high. SagUGL D115N and D175N were incapable of acting on CΔ6S even when 10 mM substrate and 0.1 mg ml^{-1} enzyme were used (Fig. 1-5d). The increase in K_m value in SagUGL D115N but not in D175N and the drastic decrease in k_{cat} in both are similar to those seen in BacillusUGL D88N and D149N.

As indicated by comparison of the active site structure, some amino acid residues interacting with substrate at subsite +1 are different between SagUGL and BacillusUGL. This difference is probably crucial for determining substrate specificity. To clarify the effects of difference in residues at subsite +1 on substrate specificity of UGLs, SagUGL Arg-236 was replaced by His or Ala. Similar to the case of SagUGL D115N, kinetics studies of SagUGL R236H and R236A were difficult because of their higher K_m values. Unlike SagUGL D115N and D175N, both SagUGL R236H and R236A effectively degraded CΔ6S (Fig. 1-5d). As shown in TLC using high concentration of substrate and enzyme (Fig. 1-5e), SagUGL R236H and R236A were almost inactive on CΔ4S, although the substrate was degraded by the wild-type enzyme. These results indicate that SagUGL Arg-236 is one of the residues involved in its activity for the sulfated substrate. This is supported by the energy-minimized model of SagUGL binding to CΔ4S (Fig. 1-6d). Arg-236 and Gln-237 can bind to the sulfate group of CΔ4S through the formation of hydrogen bonds.

In conclusion, in response to GAGs, pathogenic streptococci produce UGL together with polysaccharide lyase responsible for completely degrading GAGs, and the streptococcal UGLs prefer sulfated substrates. This suggests that the substrate specificity is feasible for bacterial infection through degradation of mammalian extracellular matrices with sulfate groups. The amino acid sequence in the active site specific to streptococcal UGLs is probably involved in interacting with the substrate sulfate group.

References

1. Glaser, P., Rusniok, C., Buchrieser, C., Chevalier, F., Frangeul, L., Msadek, T., Zouine, M., Couvé, E., Lalioui, L., Poyart, C., Trieu-Cuot, P., and Kunst, F. (2002) Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Mol. Microbiol.* **45**, 1499-1513
2. Hoskins, J., Alborn, W. E., Arnold, J., Blaszczak, L. C., Burgett, S., DeHoff, B. S., Estrem, S. T., Fritz, L., Fu, D. J., Fuller, W., Geringer, C., Gilmour, R., Glass, J. S., Khoja, H., Kraft, A. R., Lagace, R. E., LeBlanc, D. J., Lee, L. N., Lefkowitz, E. J., Lu, J., Matsushima, P., McAhren, S. M., McHenney, M., McLeaster, K., Mundy, C. W., Nicas, T. I., Norris, F. H., O'Gara, M., Peery, R. B., Robertson, G. T., Rockey, P., Sun, P. M., Winkler, M. E., Yang, Y., Young-Bellido, M., Zhao, G. S., Zook, C. A., Baltz, R. H., Jaskunas, S. R., Rosteck, P. R., Skatrud, P. L., and Glass, J. I. (2001) Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J. Bacteriol.* **183**, 5709-5717
3. Ferretti, J. J., McShan, W. M., Ajdic, D., Savic, D. J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, A. N., Kenton, S., Lai, H. S., Lin, S. P., Qian, Y. D., Jia, H. G., Najar, F. Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X. L., Clifton, S. W., Roe, B. A., and McLaughlin, R. (2001) Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci. USA* **98**, 4658-4663
4. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
5. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
6. Hashimoto, W., Kobayashi, E., Nankai, H., Sato, N., Miya, T., Kawai, S., and Murata, K. (1999) Unsaturated glucuronyl hydrolase of *Bacillus* sp. GL1: novel enzyme prerequisite for metabolism of unsaturated oligosaccharides produced by polysaccharide lyases. *Arch. Biochem. Biophys.* **368**, 367-374
7. Bradford, M. M. (1976) A rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254
8. Laemmli, U. K. (1970) Cleavage of structural proteins during assembly of head of bacteriophage T4. *Nature* **227**, 680-685
9. Hashimoto, W., Maesaka, K., Sato, N., Kimura, S., Yamamoto, K., Kumagai, H., and Murata, K. (1997) Microbial system for polysaccharide depolymerization: enzymatic route for gellan depolymerization by *Bacillus* sp. GL1. *Arch. Biochem. Biophys.* **339**, 17-23
10. Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264
11. Hashimoto, W., Suzuki, H., Yamamoto, K., and Kumagai, H. (1995) Effect of site-directed mutations on processing and activity of γ-glutamyltranspeptidase of *Escherichia coli* K-12. *J. Biochem.* **118**, 75-80
12. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326
13. Vagin, A. A., and Isupov, M. N. (2001) Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1451-1456
14. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the *CCP4* suite and current developments. *Acta Crystallogr. D. Biol.*

Crystallogr. **67**, 235-242
15. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* **53**, 240-255
16. Emsley, P., and Cowtan, K. (2004) *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126-2132
17. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *PROCHECK* - a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291
18. DeLano, W. L. (2004) *The PyMOL Molecular Graphics System*. DeLano Scientific LLC, San Carlos, CA
19. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The protein data bank. *Nucleic Acids Res.* **28**, 235-242
20. Mori, S., Akao, S., Nankai, H., Hashimoto, W., Mikami, B., and Murata, K. (2003) A novel member of glycoside hydrolase family 88: overexpression, purification, and characterization of unsaturated β-glucuronyl hydrolase of *Bacillus* sp. GL1. *Protein Expr. Purif.* **29**, 77-84
21. Gase, K., Ozegowski, J., and Malke, H. (1998) The *Streptococcus agalactiae hylB* gene encoding hyaluronate lyase: completion of the sequence and expression analysis. *Biochim. Biophys. Acta Gene Struct. Expr.* **1398**, 86-98
22. Berry, A. M., Lock, R. A., Thomas, S. M., Rajan, D. P., Hansman, D., and Paton, J. C. (1994) Cloning and nucleotide sequence of the *Streptococcus pneumoniae* hyaluronidase gene and purification of the enzyme from recombinant *Escherichia coli*. *Infect. Immun.* **62**, 1101-1108
23. Hynes, W. L., Dixon, A. R., Walton, S. L., and Aridgides, L. J. (2000) The extracellular hyaluronidase gene (*hylA*) of *Streptococcus pyogenes*. *FEMS Microbiol. Lett.* **184**, 109-112
24. Brinkkotter, A., Kloss, H., Alpert, C. A., and Lengeler, J. W. (2000) Pathways for the utilization of *N*-acetyl-galactosamine and galactosamine in *Escherichia coli*. *Mol. Microbiol.* **37**, 125-135
25. Yum, D. Y., Lee, B. Y., and Pan, J. G. (1999) Identification of the *yqhE* and *yafB* genes encoding two 2,5-diketo-D-gluconate reductases in *Escherichia coli*. *Appl. Environ. Microbiol.* **65**, 3341-3346
26. Myette, J. R., Shriver, Z., Kiziltepe, T., McLean, M. W., Venkataraman, G., and Sasisekharan, R. (2002) Molecular cloning of the heparin/heparan sulfate Δ4,5 unsaturated glycuronidase from *Flavobacterium heparinum*, its recombinant expression in *Escherichia coli*, and biochemical determination of its unique substrate specificity. *Biochemistry* **41**, 7424-7434
27. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) SCOP - a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536-540
28. Juy, M., Amit, A. G., Alzari, P. M., Poljak, R. J., Claeyssens, M., Beguin, P., and Aubert, J. P. (1992) Three-dimensional structure of a thermostable bacterial cellulase. *Nature* **357**, 89-91
29. Itoh, T., Mikami, B., Maru, I., Ohta, Y., Hashimoto, W., and Murata, K. (2000) Crystal structure of *N*-acyl-D-glucosamine 2-epimerase from porcine kidney at 2.0 Å resolution. *J. Mol. Biol.* **303**, 733-744
30. Itoh, T., Ochiai, A., Mikami, B., Hashimoto, W., and Murata, K. (2006) A novel glycoside hydrolase family 105: the structure of family 105 unsaturated rhamnogalacturonyl hydrolase complexed with a disaccharide in comparison with family 88 enzyme complexed with the disaccharide. *J. Mol. Biol.* **360**, 573-585

31. Itoh, T., Akao, S., Hashimoto, W., Mikami, B., and Murata, K. (2004) Crystal structure of unsaturated glucuronyl hydrolase, responsible for the degradation of glycosaminoglycan, from *Bacillus* sp. GL1 at 1.8 Å resolution. *J. Biol. Chem.* **279**, 31804-31812
32. Itoh, T., Hashimoto, W., Mikami, B., and Murata, K. (2006) Crystal structure of unsaturated glucuronyl hydrolase complexed with substrate. *J. Biol. Chem.* **281**, 29807-29816
33. Itoh, T., Hashimoto, W., Mikami, B., and Murata, K. (2006) Substrate recognition by unsaturated glucuronyl hydrolase from *Bacillus* sp. GL1. *Biochem. Biophys. Res. Commun.* **344**, 253-262

CHAPTER II

Structural determinants in streptococcal unsaturated glucuronyl hydrolase for recognition of sulfated substrates

Genes for UGL are present in the genome of pathogenic and indigenous bacteria, such as streptococci, enterococci, lactobacilli, mycoplasma, bacteroides, and vibrios. As described in CHAPTER I, the enzyme gene is inducibly transcribed in *S. agalactiae* cells grown in the presence of GAG, indicating that the bacterium produces UGL as well as polysaccharide lyase for complete degradation of GAGs. SagUGL, SpnUGL, and SpyUGL degrade unsaturated chondroitin and hyaluronan disaccharides, and exhibit a preference for sulfated chondroitin disaccharides, especially for CΔ6S. On the other hand, UGL of *Bacillus* sp. GL1, a soil bacterium, prefers unsulfated chondroitin disaccharide. Although the binding site of an amino sugar is suggested to cause the differing substrate specificity between streptococcal and bacillus UGLs, this hypothesis remains to be experimentally confirmed.

This CHAPTER deals with structural determinants in streptococcal UGL for recognition of sulfated substrates through X-ray crystallography, kinetics of site-directed mutants, and molecular conversion of bacterial UGLs with altered substrate specificity.

Materials and Methods

Materials. Unsaturated GAG disaccharides were purchased from Seikagaku Biobusiness or Sigma-Aldrich. Silica gel 60/Kieselguhr F_{254} TLC plates were obtained from Merck. Restriction endonucleases and DNA-modifying enzymes were from Toyobo. Other analytical grade chemicals were obtained from commercial sources.

Microorganisms, Culture Conditions, and Purification. For expression of SagUGL and BacillusUGL, *E. coli* strain HMS174(DE3) cells transformed with pET21b-SagUGL and *E. coli* strain BL21(DE3) cells with pET3a-BacillusUGL (1) were cultivated as described in CHAPTER I. *E. coli* cells harboring pET21b-SagUGL or pET3a-BacillusUGL were grown in LB medium (2), collected by centrifugation at 6,700 × *g* and 4°C for 5 min, and resuspended in 20 mM KPB (pH 7.0). The cells were ultrasonically disrupted (Insonator model 201M) at 9 kHz and 0°C for 5 min, and the clear solution obtained by centrifugation at 28,000 × *g* and 4°C for 20 min was used as a cell extract. SagUGL and BacillusUGL were purified from the cell extract to homogeneity by several steps of column chromatography. Briefly, SagUGL and BacillusUGL were purified by anion

exchange chromatography (TOYOPEARL DEAE-650M), followed by gel filtration chromatography (HiLoad 16/60 Superdex 75 pg), and finally by anion exchange chromatography (monoQ 10/100 GL). The degree of purification was checked by SDS-PAGE (3).

TLC. The products derived from unsaturated GAG disaccharides through the reaction of bacterial UGLs were separated by TLC using a solvent system of 1-butanol/acetic acid/water (3/2/2, v/v). The following disaccharides were used: ΔHA; CΔ0S; CΔ4S; sulfate group-free unsaturated heparin disaccharide, ΔGlcUA-1,4-GlcNAc (HΔNAc0S); unsaturated heparin disaccharide with a sulfate group at the *N* position of GlcN, ΔGlcUA-1,4-GlcNS (HΔNS); unsaturated heparin disaccharide with a sulfate group at the C-6 position of GlcN, ΔGlcUA-1,4-GlcN6S (HΔ6S); unsaturated heparin disaccharide with a sulfate group at the C-6 position of GlcNAc, ΔGlcUA-1,4-GlcNAc6S (HΔNAc6S); unsaturated heparin disaccharide with sulfate groups at *N* and C-6 positions of GlcN, ΔGlcUA-1,4-GlcNS6S (HΔNS6S); unsaturated heparin disaccharide with sulfate groups at the C-2 position of ΔGlcUA and the *N* position of GlcN, ΔGlcUA2S-1,4-GlcNS (HΔ2'SNS); unsaturated heparin disaccharide with sulfate groups at the C-2 position of ΔGlcUA and the *N* and C-6 positions of GlcN, ΔGlcUA2S-1,4-GlcNS6S (HΔ2'SNS6S). The products were visualized by heating the TLC plates at 130°C for 5 min after spraying with 10% sulfuric acid in ethanol.

Crystallization and X-ray Diffraction. Purified SagUGL enzymes of wild-type and mutant D175N with Asp-175 replaced by Asn were concentrated by ultrafiltration using Centriprep (10,000 molecular weight cut-off) to 10 mg ml^{-1} and 5 mg ml^{-1}, respectively. Both wild-type enzyme and D175N were crystallized by sitting-drop vapor diffusion. The 3 μl of proteins were mixed with equal volume of a reservoir solution. The reservoir solution for wild-type enzyme crystallization contains 30% polyethylene glycol 200, 1% polyethylene glycol 3,000, and 0.1 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES, pH 7.0). Wild-type crystals grew up at 20°C for a week. The reservoir solution for D175N crystallization includes 40% ethylene glycol, 5% polyethylene glycol 3,000, and 0.1 M HEPES (pH 7.5). D175N was crystallized at 20°C for a month. In order to prepare a complex form of SagUGL and CΔ6S, the D175N crystal was soaked at 20°C for 10 min in a reservoir solution containing 0.2 M CΔ6S before X-ray diffraction experiments. Crystals were placed in a cold nitrogen gas stream at –173°C. X-ray diffraction images of crystals were collected using a Jupiter 210 CCD detector for the wild-type crystal or a Quantum 210 CCD detector (ADSC) for D175N crystals with synchrotron radiation at a wavelength of 1.0000 Å at the BL38B1 station of SPring-8. The data were processed and scaled with the *HKL2000*

program (4).

Structure Determination and Refinement. The structure was determined through molecular replacement with the *Molrep* program (5) supplied in *CCP4* program package (6) by using determined coordinates of SagUGL wild-type in CHAPTER I (PDB ID: 2ZZR) as an initial model. Structure refinement was conducted with the *Refmac5* program (7). Randomly selected 5% reflections were excluded from refinement and used to calculate R_{free}. After each refinement cycle, the model was adjusted manually using the *winCoot* program (8). Water molecules were incorporated where the difference in density exceeded 3.0 σ above the mean and the $2F_o–F_c$ map showed a density of more than 1.0 σ. The structure of the enzyme–sugar complex was refined using the *winCoot* program with the chondroitin disaccharide parameter file constructed at the PRODRG site (http://davapc1.bioch.dundee.ac.uk). Protein models were superimposed, and their r.m.s.d. was determined with the *LSQKAB* program (9), a part of the *CCP4* program package. Final model quality was checked with the *PROCHECK* program (10). Figures for protein structures were prepared using the *PyMol* program (11). Electric charge on the molecular surface of bacterial UGLs was calculated using the *APBS* program (12). Coordinates used in this CHAPTER were taken from PDB (13). The atomic coordinates and structure factors of SagUGL wild-type (ID: 3ANJ), D175N mutant (ID: 3ANI), and D175N mutant complexed with CΔ6S (ID: 3ANK) have been deposited in the RCSB PDB (http://www.rcsb.org/).

Site-directed Mutagenesis. Thr-235, Ser-365, Ser-368, and Lys-370 of SagUGL were substituted with Ala, His, Gly, and Ile, respectively, and His-210, His-339, Gly-342, and Ile-344 of BacillusUGL were with Arg, Ser, Ser, and Lys, respectively. UGL mutants were constructed using a QuikChange site-directed mutagenesis kit. The plasmid pET21b-SagUGL or pET3a-BacillusUGL (1) was used as a PCR template, and the synthetic oligonucleotides were used as sense and antisense primers shown in Table 2-1. PCR was carried out using KOD-FX polymerase (Toyobo) in place of Pfu polymerase. Mutations were confirmed by dideoxy-chain termination method (14) using automated DNA sequencer model 3730xl. The cells of *E. coli* host strain [HMS174(DE3)] were transformed with the mutant plasmids. Expression and purification of the mutants were conducted by using the same procedures as for SagUGL or BacillusUGL wild-type as described above. DNA manipulations such as plasmid isolation, subcloning, transformation, and gel electrophoresis were performed as described (2).

Kinetic Analysis. Kinetics parameters of SagUGLs (wild-type, T235A, S365H, S368G, K370I) and BacillusUGLs (wild-type, H210R, H339S, G342S, and I344K) toward CΔ6S or CΔ0S were

determined as follows: the activity of UGLs was assayed at 30°C by monitoring the decrease in absorbance at 235 nm arising from the double bond (molar extinction coefficient ε_{235} = 4,800 M^{-1} cm^{-1}) in the substrates. The reaction mixtures consisted of substrate (The range of substrate concentration is fixed 0.05–1.0 mM because the absorbance at 235 nm of the substrate at over 1 mM exceeds measurement limitations on the spectrometer), 20 mM Tris–HCl (pH 7.5), and enzyme. K_m and k_{cat} were calculated using the Michaelis–Menten equation with KaleidaGraph software (Synergy Software).

Table 2-1. Primers for site-directed mutagenesis

SagUGL		
T235A	sense	5′-CCGTTAAAAGGTGTCGCACGACAGGGTTATAG-3′
	antisense	5′-CTATAACCCTGTCGTGCGACACCTTTTAACGG-3′
S365H	sense	5′-CTATTGCACGGTGTGTATCATTGGCATTCAGGTAAAGGAG-3′
	antisense	5′-CTCCTTTACCTGAATGCCAATGATACACACCGTGCAATAG-3′
S368G	sense	5′-CGGTGTGTATTCGTGGCATGGAGGTAAAGGAGTAGATGAAG-3′
	antisense	5′-CTTCATCTACTCCTTTACCTCCATGCCACGAATACACACCG-3′
K370I	sense	5′-CGTGGCATTCAGGTATAGGAGTAGATGAAGGTAATATTTG-3′
	antisense	5′-CAAATATTACCTTCATCTACTCCTATACCTGAATGCCACG-3′
BacillusUGL		
H210R	sense	5′-CATCCGCGGCGGCACGCGCCAGGGCAACACCGACGGC-3′
	antisense	5′-GCCGTCGGTGTTGCCCTGGCGCGTGCCGCCGCGGATG-3′
H339S	sense	5′-CTTCATCCGCCGCGGCTCCTATTCTGTGCGCGGCGGCATC-3′
	antisense	5′-GATGCCGCCGCGCACAGAATAGGAGCCGCGGCGGATGAAG-3′
G342S	sense	5′-CTATCACGTGCGCAGTGGCATCTCGCCCGACGACTACAC-3′
	antisense	5′-GTGTAGTCGTCGGGCGAGATGCCACTGCGCACGTGATAG-3′
I344K	sense	5′-CTATCACGTGCGCGGCGGCAAATCGCCCGACGACTACACG-3′
	antisense	5′-CGTGTAGTCGTCGGGCGATTTGCCGCCGCGCACGTGATAG-3′

Underline indicates the mutation site.

Results and Discussion

Streptococcal UGL Involved in Heparin Degradation

In the previous report and CHAPTER I (15), bacillus and streptococcal UGLs have been identified to degrade unsaturated chondroitin and hyaluronan disaccharides. Both chondroitin and hyaluronan include β-GlcUA as a component, whereas α-IdoUA, a C-5 epimer of GlcUA, is predominant (the ratio in uronates, >70%) in heparin molecules (16). There is also a difference in glycosidic bond pattern between chondroitin/hyaluronan and heparin/heparan sulfate. The 1,3-glycosidic bond is present between uronate and amino sugar residues in chondroitin and hyaluronan, although both heparin and heparan sulfate contain the 1,4-glycosidic bond between

uronate and amino sugar residues. Thus, the enzyme activity of SagUGL was investigated using unsaturated heparin disaccharides with and without sulfate group(s) as a substrate. HΔNAc0S was completely degraded to unsaturated uronic acid and GlcNAc by SagUGL (Fig. 2-1a, lane 5). In combination with previously reported data, this result indicates that the enzyme acts on unsaturated α-IdoUA (ΔIdoUA) as well as unsaturated β-GlcUA residues in substrates and cleaves both glycosidic bonds 1,3 and 1,4. The formation of a double bond between C-4 and C-5 atoms leads to loss of epimerization in GlcUA and IdoUA. In fact, C-3, C-4, C-5, and C-6 atoms of ΔGlcUA were determined to be located in a single plane through structural analysis of UGL and unsaturated GAG disaccharide (17). Thus, the degradation of both unsaturated chondroitin and heparin disaccharides by UGL was chemically appropriate because each nonreducing terminus showed the same conformation.

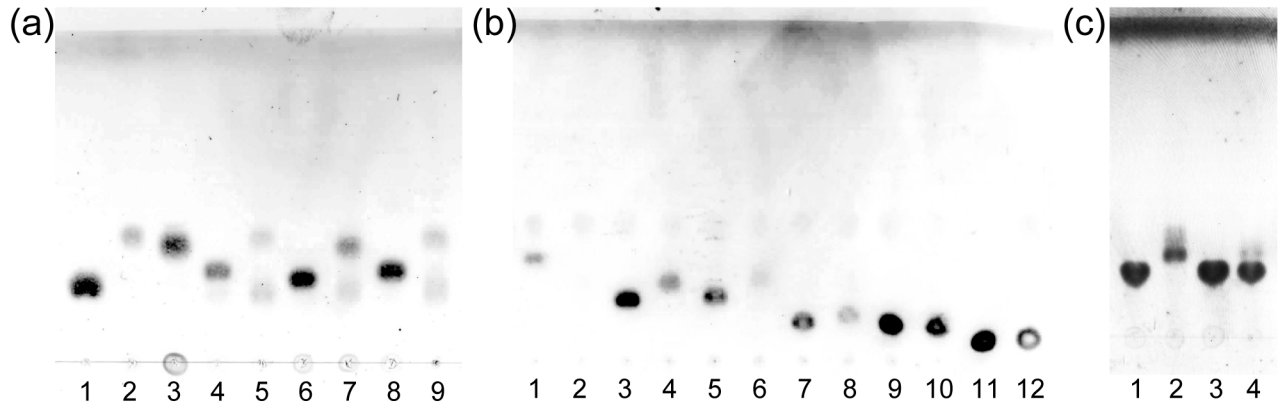

Figure 2-1. UGL reaction analysis by TLC. (a) Unsaturated GAG disaccharides were incubated at 30°C with 5 mg ml^{-1} SagUGL and resultant products were detected on the TLC plate. *Lane 1*, GlcUA; *lane 2*, GlcNAc; *lane 3*, GalNAc; *lane 4*, HΔNAc0S; *lane 5*, HΔNAc0S with SagUGL; *lane 6*, CΔ0S; *lane 7*, CΔ0S with SagUGL; *lane 8*, ΔHA; and *lane 9*, ΔHA with SagUGL. (b) Unsaturated heparin disaccharides were incubated at 30°C with 5 mg ml^{-1} SagUGL and resultant products detected on the TLC plate. *Lane 1*, HΔNAc0S; *lane 2*, HΔNAc0S incubated with SagUGL; *lane 3*, HΔNS; *lane 4*, HΔNS incubated with SagUGL; *lane 5*, HΔNAc6S; *lane 6*, HΔNAc6S incubated with SagUGL; *lane 7*, HΔNS6S; *lane 8*, HΔNS6S incubated with SagUGL; *lane 9*, HΔ2'SNS; *lane 10*, HΔ2'SNS incubated with SagUGL; *lane 11*, HΔ2'SNS6S, *lane 12*, HΔ2'SNS6S incubated with SagUGL. (c) CΔ4S was incubated at 30°C with 1 mg ml^{-1} SagUGL or 1 mg ml^{-1} BacillusUGLs for 8 h and resultant products were detected on the TLC plate. *Lane 1*, CΔ4S; *lane 2*, CΔ4S with SagUGL; *lane 3*, CΔ4S with BacillusUGL wild-type; and *lane 4*, CΔ4S with BacillusUGL H210R.

The activity of UGL on ΔGlcUA- and ΔIdoUA-containing GAG disaccharides suggests that each GAG in mammalian extracellular matrices is first depolymerized to unsaturated disaccharides by a specific polysaccharide lyase, and the resultant unsaturated disaccharides are degraded to the

constituent monosaccharides by a single enzyme, UGL. Sulfate-bound disaccharides, such as unsaturated heparin disaccharides with sulfate group(s) at either or both the C-6 position of the GlcNAc residue or the nitrogen position of the GlcN residue, were also degraded by SagUGL (Fig. 2-1b). Similar to chondroitin and heparin, heparan sulfate and dermatan sulfate include IdoUA and/or GlcUA and sulfate group(s) in their molecules (16). Therefore, streptococcal UGLs are considered as one of the key enzymes for degradation of all uronate-including GAGs, chondroitin, hyaluronan, heparin, heparan sulfate, and dermatan sulfate after reactions of polysaccharide lyases.

Enzyme Activity for Sulfated Substrates

Streptococcal UGLs, including SagUGL, showed a preference for sulfate group-bound unsaturated disaccharides from GAGs (*e.g.* chondroitin and heparin). Because this preference of streptococcal UGLs was thought to be dependent on substrate binding, catalytic action, or both, SagUGL and BacillusUGL were kinetically analyzed using sulfate group-free (CΔ0S) and -bound (CΔ6S) unsaturated chondroitin disaccharides as substrates (Table 2-2). Kinetic parameters listed in Table 2-2 were interpreted as follows. Kinetic parameters with higher Michaelis constants (K_m over 1 mM) were treated as estimations, because kinetic studies of bacterial UGLs with the higher concentrations (over 1 mM) of substrate were difficult to perform because of their high absorbance at 235 nm. Kinetic parameters of SagUGL toward CΔ0S and CΔ6S were determined. The K_m value (1.3 mM) toward CΔ0S was over 10-fold higher than that (0.10 mM) toward CΔ6S, whereas turnover number (k_{cat}) toward CΔ0S was about 3.8-fold smaller than that toward CΔ6S. In comparison with the great difference in the affinity for substrate (K_m), the difference in k_{cat} was considered to be insignificant. This result indicates that the enzyme exhibits a higher affinity for CΔ6S than CΔ0S, and its substrate specificity mostly depends on substrate binding. In contrast, BacillusUGL showed a preference for the unsulfated substrate rather than the sulfated substrate because of its high affinity (around 50-fold), although the difference in k_{cat} of the enzyme toward CΔ0S and CΔ6S were less than 4-fold.

Kinetics of SagUGL and BacillusUGL demonstrated that the activity on the sulfated substrate was mainly dependent on the affinity for the substrate rather than turnover number. This lesser effect of k_{cat} on substrate specificity was likely due to the multiple actions of UGL strictly on ΔGlcUA but not on amino sugar (17), *i.e.* (i) proton donation to the C-4 atom of ΔGlcUA, (ii) deprotonation of the water molecule, (iii) addition of water molecule to the C-5 atom of ΔGlcUA, (iv) cleavage of the glycosidic bond, and (v) a conversion of ΔGlcUA to α-keto acid. Hereafter, substrate specificity was analyzed based on the affinity (K_m) for the substrate.

Table 2-2. Kinetic parameters of bacterial UGLs

	CΔ6S			CΔ0S		
	K_m (mM)	k_{cat} (s^{-1})	$k_{cat}K_m^{-1}$ (s^{-1} mM^{-1})	K_m (mM)	k_{cat} (s^{-1})	$k_{cat}K_m^{-1}$ (s^{-1} mM^{-1})
SagUGL						
wild-type	0.10 ± 0.03	10 ± 1	100	1.3 ± 0.1	2.7 ± 0.1	2.1
T235A	1.8 ± 0.3	4.0 ± 0.6	2.3	ND[a]	ND	ND
S365H	2.4 ± 0.2	3.9 ± 0.2	1.6	0.76 ± 0.22	1.7 ± 0.3	2.2
S368G	0.19 ± 0.03	25 ± 1	130	1.4 ± 0.3	11 ± 2	7.7
K370I	1.5 ± 0.2	42 ± 4	28	0.37 ± 0.07	4.0 ± 0.3	11
BacillusUGL						
wild-type	19 ± 7	55 ± 18	3.0	0.38 ± 0.04	14 ± 1	37
H339S	7.2 ± 1.5	25 ± 5	3.4	0.86 ± 0.15	17 ± 1	20
G342S	7.0 ± 2.7	21 ± 7	3.0	0.50 ± 0.04	18 ± 1	36
I344K	2.2 ± 0.1	40 ± 1	18	0.57 ± 0.09	24 ± 2	42

[a] ND, kinetic parameters could not be determined due to the very low activity.

The substrate specificity of SagUGL is suggestive of its structural features for specific binding to sulfate groups in the substrate. To clarify structural determinants for sulfate binding, X-ray crystallography of the enzyme–substrate complex was performed.

Loop Movement over the Active Cleft

In CHAPTER I, SagUGL was crystallized in a drop solution using ammonium sulfate as the major precipitant, and the crystal structure of the ligand-free enzyme was determined at a 1.75 Å resolution. Many experiments were conducted to prepare the enzyme–substrate complex through a soaking treatment of crystals in the substrate solution, but the complex was failed to obtain. Thus, crystallization conditions were rescreened to accommodate the substrate at the active site of the enzyme. Another crystal of SagUGL wild-type formed in a drop solution using polyethylene glycol as the major precipitant, and it showed crystallographic properties different from the crystal described in CHAPTER I. This crystal belongs to the *C2* space group with unit cell dimensions of a = 104.8 Å, b = 53.2 Å, c = 70.1 Å, and β =96.6°. Furthermore, the crystal structure was determined at 1.95 Å resolution via molecular replacement using the structure determined in CHAPTER I (PDB ID: 2ZZR) as an initial model. The details of data collection and model refinement statistics are summarized in Table 2-3.

The overall structure, α_6/α_6-barrel, of SagUGL in the *C*2 crystal was very similar to the determined structure (CHAPTER I) (Fig. 2-2a), although residues 151–171 were not assigned due to disorder. Furthermore, a conformational change was observed, *i.e.* loop movement over the active cleft, in this model that suggests that the loop (residues 219–236) has flexible motion. Because this

Table 2-3. Data collection and refinement statistics

	wild-type	D175N	D175N–CΔ6S
Data collection			
Wavelength (Å)	1.0000	1.0000	1.0000
Space group	*C*2	*C*2	*C*2
Unit cell parameters (Å, deg)	a = 104.8, b = 53.2, c = 70.1, β = 96.6	a = 105.1, b = 53.4, c = 70.2, β = 96.5	a = 105.2, b = 53.3, c = 70.1, β = 96.6
Resolution limit (Å)	50.00–1.95 (2.02–1.95)[a]	30.00–2.50 (2.59–2.50)	20.00–2.02 (2.09–2.02)
Measured reflections	103,030	50,235	96,246
Unique reflections	28,120	13,575	25,199
Redundancy	3.7 (3.1)	3.7 (3.8)	3.8 (3.8)
Completeness (%)	99.0 (91.9)	99.0 (98.4)	98.4 (97.7)
I/σ (I)	16.9 (2.62)	21.5 (4.90)	19.3 (4.30)
R_{merge}	0.060 (0.29)	0.064 (0.35)	0.071 (0.38)
Refinement			
R_{work}	0.189	0.187	0.174
R_{free}	0.213	0.260	0.209
No. of molecules/asymmetric unit	1	1	1
No. of nonhydrogen atoms			
Protein	3,250	3,218	3,206
CΔ6S	0	0	30
Water molecules	109	61	160
Average isotropic *B*-factor (Å^2)			
Protein	25.9	35.8	21.9
CΔ6S			42.6
Water molecules	28.7	31.6	33.5
r.m.s.d. from ideal			
Bond lengths (Å)	0.007	0.008	0.007
Bond angle (deg)	0.957	1.002	0.997
Ramachandran plot (%)			
Favored region	91.4	91.2	90.2
Allowed region	8.6	8.8	9.8
Outlier region	0	0	0

[a] Data in highest resolution shells are given in parentheses.

loop covers the active cleft to interact with the substrate in the resulting structure of the enzyme–substrate complex (Fig. 2-2c), the enzyme adopts a "closed form." The oscillation length of the loop at the edge (Thr-227) was estimated to be 16.3 Å (Fig. 2-2a). The loop flexibility was also evidenced based on the average *B*-factor. The *B*-factor of the loop showed a high score (40.3 Å^2), although the entire molecule of the enzyme had a *B*-factor of 21.9 Å^2. The flexibility of the loop (residues 219–236) was found to be important for enzyme activity through the following structure determination of the enzyme–substrate complex and site-directed mutagenesis

experiments. Briefly, Gly-233 and Thr-235, both of which are located in the loop region, directly interact with the substrate (Fig. 2-2c), and the T235A mutant exhibited little enzyme activity (Table 2-2).

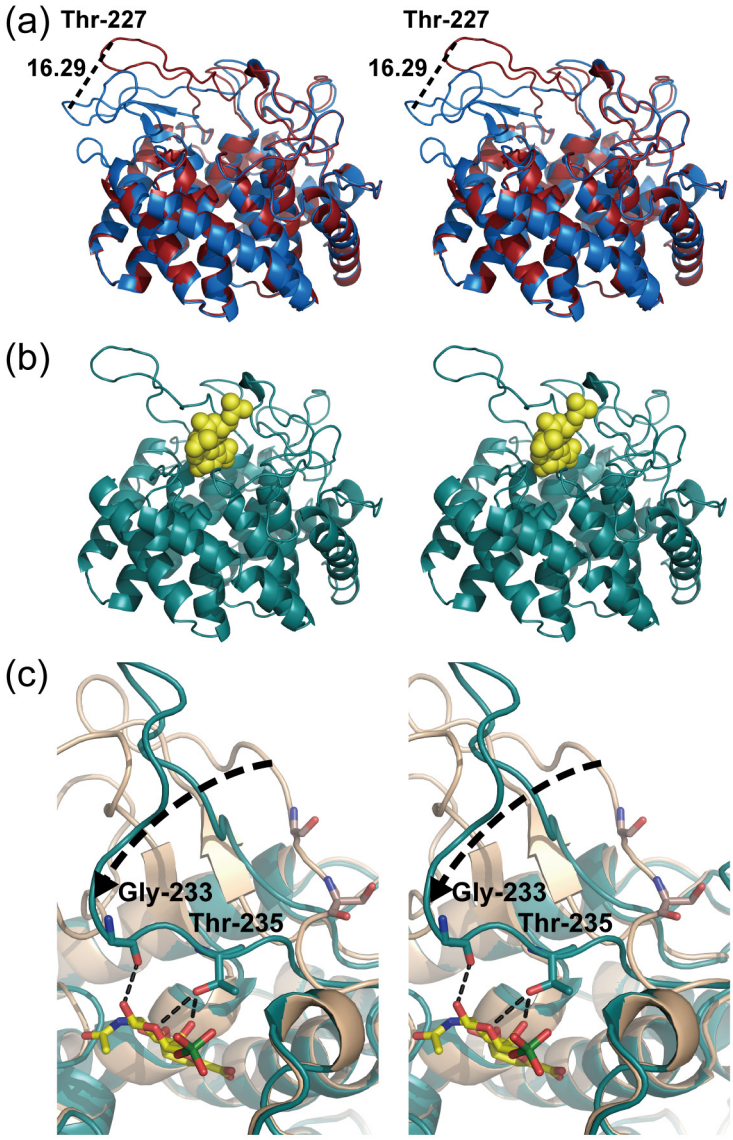

Figure 2-2. Conformational change of SagUGL (stereo-diagram). (a) Superimposition of closed (*red*) and open (*blue*) forms of SagUGL wild-type. The flexible loop moves with an oscillation width of 16.29 Å, as indicated by the *broken line*. (b) CΔ6S (*yellow*) is accommodated at the active site of SagUGL D175N (*cyan*) in closed form. (c) Gly-233 and Thr-235 binding to the substrate.

Structure of Enzyme–Substrate Complex

To analyze the interaction between SagUGL and substrate, the mutant D175N constructed in CHAPTER I was used for crystallization because the mutant exhibits no enzyme activity. The role

of Asp-175 is as follows. Asp-175 acts as a general acid catalyst and donates a proton to the double bond (C-4 atom). Subsequently, Asp-175 also acts as a general base catalyst and deprotonates the water molecule. The freshly prepared D175N crystals belonging to the *C*2 space group with and without CΔ6S were subjected to structure determination. Crystal structures of CΔ6S-free and -bound D175N were determined by the molecular replacement method using the previous structure (PDB ID: 2ZZR) as an initial model (Fig. 2-2b). The details of data collection and the model refinement statistics are summarized in Table 2-3. Both D175N and D175N–CΔ6S models are structurally identical to the closed form of wild-type (D175N, r.m.s.d. = 0.71 Å for 364 C^{α}; D175N–CΔ6S, r.m.s.d. = 0.64 Å for 370 C^{α}), and lacked 20 residues (D175N, residues 153–171; D175N–CΔ6S, residues 152–171) due to disorder. The substrate CΔ6S is well fitted in the electron density map with an average *B*-factor of 42.6 $Å^2$ (Fig. 2-3a) and bound to the active cleft (Fig. 2-2b).

The crystal structure of D175N–CΔ6S revealed the binding mode of the substrate CΔ6S molecule to the active site cleft (Fig. 2-3b). There are several interactions between the enzyme and substrate through hydrogen bonds and van der Waals contacts (Table 2-4). Subsites are defined such that −n represents the nonreducing terminus and +n the reducing terminus, and cleavage occurs between the −1 and +1 sites (18). Because the glycosidic bond in unsaturated GAG disaccharides is cleaved through the UGL reaction, uronate and amino sugar residues are positioned at −1 and +1 subsites, respectively. ΔGlcUA residue is accommodated at the −1 subsite via six hydrogen bonds by four residues (Asp-115, Asn-175, Arg-247 and Trp-251) and stacking with six residues (Trp-69, Asp-115, Phe-118, Asn-175, Trp-245, and Arg247). In particular, a Trp-69 residue is parallel to the pyranose ring of ΔGlcUA through a stacking interaction. These interactions of SagUGL with ΔGlcUA are comparable with those of BacillusUGL with ΔGlcUA (D88N–CΔ0S) (17).

At subsite +1, seven hydrogen bonds by five residues (Asp-115, Gly-233, Thr-235, Ser-365 and Ser-368) as well as stacking and electrostatic interactions with five residues (Trp-69, Thr-235, Tyr-364, Ser-365, and Lys-370) are formed between the enzyme and the GalNAc residue with a sulfate group at the C-6 position (GalNAc6S). Distinct from those interactions between the enzyme and ΔGlcUA, there is a difference in the binding mode between SagUGL (D175N –CΔ6S) and BacillusUGL (D88N–CΔ0S) to the amino sugar residue (17), although the difference is partly generated from the presence of a sulfate group in the amino sugar (GalNAc6S). Of note, Gly-233 and Thr-235 bound to the pyranose frame of GalNAc6S are located in the flexible loop.

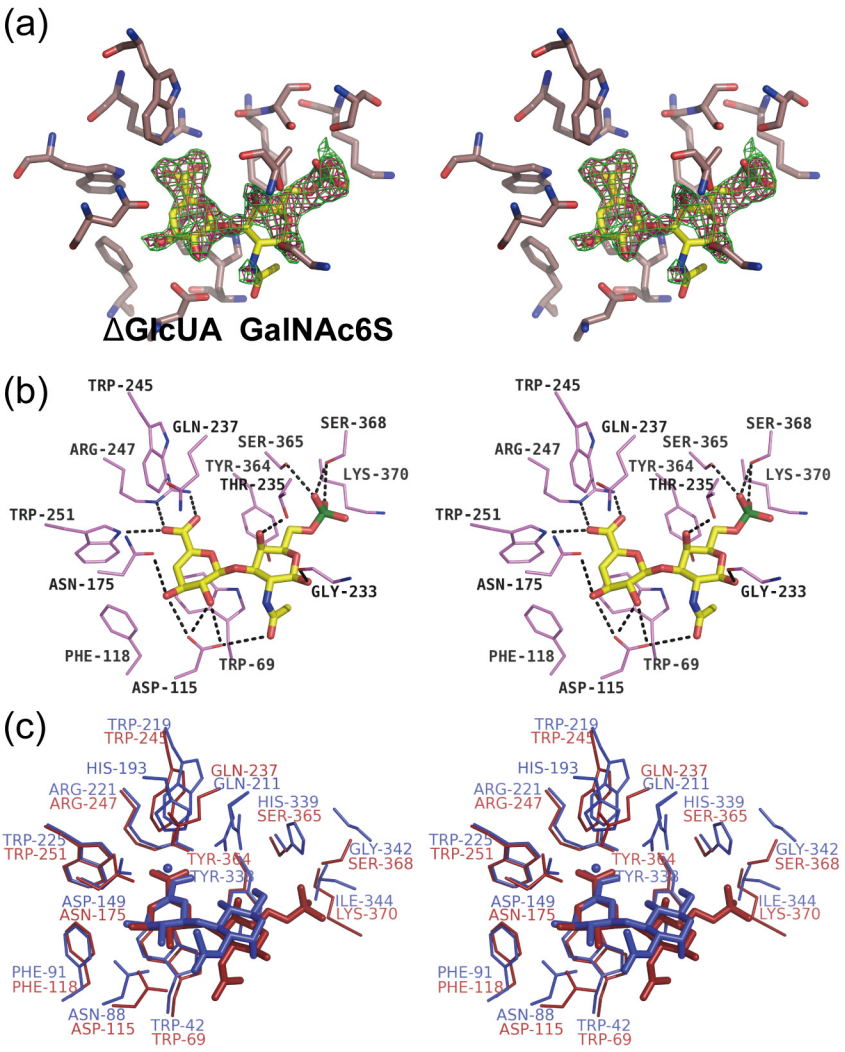

Figure 2-3. Active site structure of SagUGL D175N in complex with CΔ6S (stereo-diagram). (a) Electron density of CΔ6S in the omit (F_o−F_c) map calculated without the substrate and contoured at the 2.5 σ (*green*) and 3.0 σ (*red*) levels. (b) Interaction of SagUGL D175N with CΔ6S. Several residues bind to CΔ6S through the formation of hydrogen bonds (*broken lines*). Atoms C, O, N, and S of CΔ6S are colored by *yellow*, *pink*, *blue*, and *green*, respectively. (c) Superimposition of SagUGL D175N–CΔ6S complex (*red*) and BacillusUGL D88N–CΔ0S complex (*blue*).

The carbonyl oxygen atom of Gly-233 in the flexible loop is directly hydrogen-bonded to O-1 of GalNAc6S (Fig. 2-2c, Table 2-4). In addition, the side chain of Thr-235 in the loop also binds to GalNAc6S through formation of both hydrogen bonds and van der Waals contacts. To clarify the role of Thr-235 in the enzyme reaction, a site-directed mutant T235A was constructed and its enzymatic activity was determined (Table 2-2). In comparison with SagUGL wild-type (K_m = 0.10 mM), T235A exhibited less affinity (K_m = 1.8 mM) for the substrate CΔ6S. Furthermore, in the case of CΔ0S as the substrate, no enzyme activity of T235A was detected. These results suggest that

interaction of flexible loop with GalNAc6S in CΔ6S or GalNAc in CΔ0S is essential to the enzyme reaction.

Sulfate Group-binding Residues

Because SagUGL exhibits a maximal activity toward CΔ6S among various GAG disaccharides mostly due to its high affinity for the substrate (Table 2-2), hereafter, SagUGL recognition of the sulfate group at the C-6 position of the GalNAc residue in the substrate was examined. There are three direct hydrogen bonds between the sulfate group and the enzyme (Table 2-4): O-1S…Ser-365 Oγ (2.9 Å), O -1S…Ser-368 Oγ (2.6Å), and O -2S…Ser-368 Oγ (3.0Å). In addition, a posit ively charged residue, Lys-370, is situated around the negatively charged sulfate group in the substrate. No water molecules are included around the sulfate group in the D175N–CΔ6S complex. The residues Ser-365, Ser-368, and Lys-370 are completely conserved in UGLs of other pathogenic

Table 2-4. Interaction between SagUGL D175N and CΔ6S

Hydrogen bonds (< 3.3 Å)					van der Waals contact (C–C distance < 4.5 Å)			
Sugar	Atom	Protein	Atom	Distance	Sugar	Atom	Protein	Atom
				Å				
ΔGlcUA	O2	Asp-115	Oδ2	2.6	ΔGlcUA	C1	Trp-69	Cδ2, Cζ2, Cγ,
	O3	Asn-175	Oδ1	3.1				Cδ1, Cϵ2
		Asp-115	Oδ2	3.2		C2	Asp-115	Cγ
	O6A	Arg-247	Nη2	2.7		C3	Asn-175	Cγ
	O6B	Trp-251	Nϵ1	3.1			Asp-115	Cγ
		Arg-247	Nϵ	2.9			Trp-69	Cδ2, Cϵ3, Cζ3,
GalNAc6S	O1	Gly-233	O	2.8				Cη2, Cζ2, Cϵ2
	O4	Thr-235	Oγ1	2.5			Phe-118	Cδ1, Cϵ1
	O5	Thr-235	Oγ1	3.2		C4	Asn-175	Cγ
	O7	Asp-115	Oδ1	3.0			Trp-69	Cδ2, Cϵ3, Cζ3,
	O1S	Ser-365	Oγ	2.9				Cη2, Cζ2, Cϵ2
		Ser-368	Oγ	2.6			Phe-118	Cϵ1
	O2S	Ser-368	Oγ	3.0		C5	Trp-69	Cζ3, Cη2, Cζ2,
								Cϵ2
						C6	Trp-245	Cη2, Cζ2
							Trp-69	Cη2, Cζ2
							Arg-247	Cζ
					GalNAc6S	C3	Trp-69	Cδ1
						C4	Tyr-364	Cϵ1, Cζ
						C5	Tyr-364	Cϵ1, Cζ
						C6	Thr-235	Cγ2
							Tyr-364	Cϵ1, Cϵ2, Cζ
							Ser-365	Cβ
						C7	Trp-69	Cδ1
						C8	Trp-69	Cδ1

streptococcal species (*S. pneumoniae* and *S. pyogenes*) that are highly active on the sulfated substrate CΔ6S but not in BacillusUGL, which indicates a preference for the unsulfated substrate CΔ0S. In the case of BacillusUGL, Ser-365, Ser-368, and Lys-370 instead correspond to His-339, Gly-342, and Ile-344, respectively. Thus, three SagUGL mutants, S365H (Ser-365 to His), S368G (Ser-368 to Gly), and K370I (Lys-370 to Ile) were overexpressed in *E. coli* cells, and each of them was purified to homogeneity. Subsequently, the enzyme characteristics of these mutants were kinetically analyzed (Table 2-2). Compared with SagUGL wild-type, all three of these mutants, especially S365H and K370I, showed greatly increased K_m values for CΔ6S, indicating that the affinity of the mutants for the substrate was drastically reduced. In contrast, the binding affinity of S365H and K370I for CΔ0S increased and their K_m values toward CΔ0S were determined to be lower than those for the mutants toward CΔ6S. These results indicate that S365H and K370I are converted to a BacillusUGL-like enzyme with a preference for unsulfated substrates.

To verify that the Ser-365, Ser-368, and Lys-370 residues are responsible for binding to the sulfate group at the C-6 position of the GalNAc residue in CΔ6S, the sequence conversion of BacillusUGL to SagUGL-like enzyme were also carried out. Three mutants of BacillusUGL, H339S (His-339 to Ser), G342S (Gly-342 to Ser), and I344K (Ile-344 to Lys), were constructed and subjected to an enzyme assay. The K_m values of the three mutants toward CΔ6S were determined to be lower than that of BacillusUGL wild-type, demonstrating that the mutants showed higher affinity for CΔ6S than wild-type. To the contrary, the affinity of the mutants for CΔ0S was slightly lower than that of wild-type, although the mutants still exhibited a preference for the unsulfated substrate rather than the sulfated substrate.

Because sulfate groups are negatively charged, the electrostatic level on the molecular surface of both SagUGL and BacillusUGL were examined. Although the active pocket binding to acidic ΔGlcUA was positively charged on both enzymes, the SagUGL region that interacted with the sulfate group showed a highly positive charge mainly due to the presence of Lys-370 (Fig. 2-4). This positive charge, which is crucial for sulfate binding, was also supported by mutational analysis of K370I. In contrast, the corresponding region of BacillusUGL was determined to be negatively charged.

In CHAPTER I, Arg-236 was suggested to be one of the key residues involved in degradation of unsaturated chondroitin disaccharide with a sulfate group at the C-4 position of the GalNAc residue (CΔ4S) through site-directed mutagenesis and *in silico* modeling. Because Arg-236 corresponds to His-210 in BacillusUGL, a mutant of BacillusUGL H210R was constructed and

characterized. The substrate CΔ4S was degraded by BacillusUGL H210R similarly to SagUGL but not by BacillusUGL wild-type (Fig. 2-1c). This molecular conversion indicates that Arg-236 of SagUGL is responsible for recognition of the sulfate group at the C-4 position of the GalNAc residue in CΔ4S.

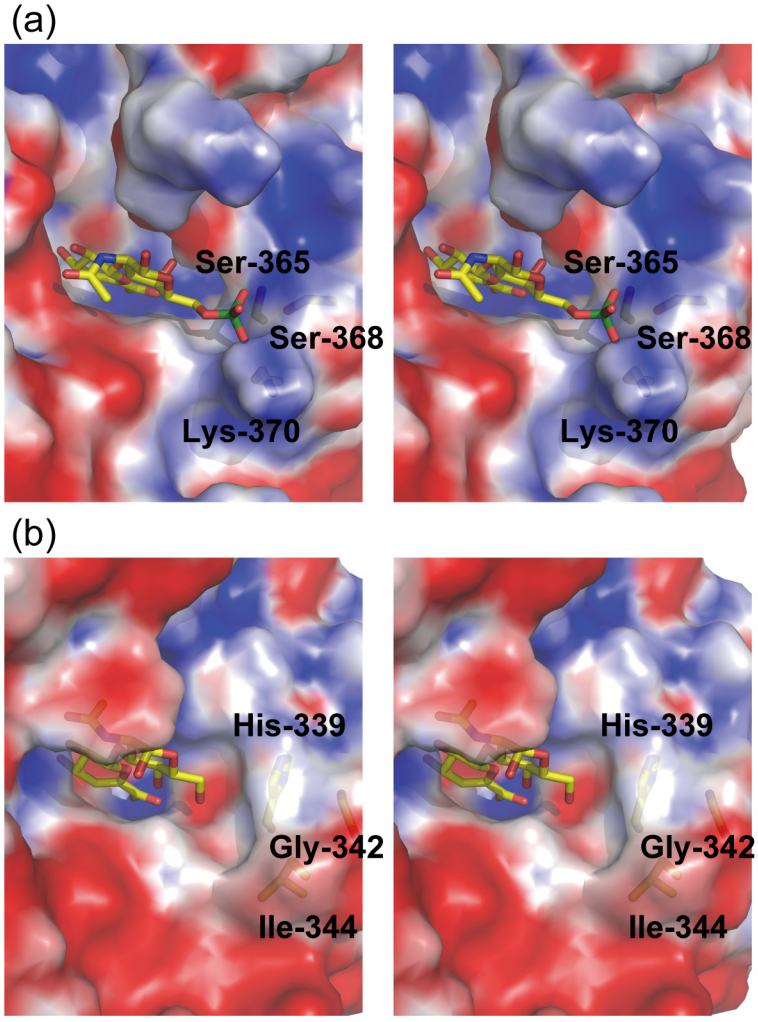

Figure 2-4. Electric charge on the molecular surface of UGL active site (stereo-diagram). (a) SagUGL D175N–CΔ6S complex. (b) BacillusUGL D88N–CΔ0S complex. Positive and negative charges at pH 7.0 are colored *blue* and *red*, respectively.

In this CHAPTER, SagUGL was found to act on unsaturated chondroitin, hyaluronan, and heparin disaccharides, demonstrating that this enzyme is involved in complete degradation of various uronate-containing GAGs after treatment of polysaccharide lyases. Other than bacillus and streptococcal enzymes, previous studies have also characterized two UGLs of a soil isolate, *P. heparinus* (19,20). One of the *Pedobacter* UGLs was found to be specific for unsaturated heparin

disaccharides with a 1,4-glycosidic bond, and the other was determined to prefer unsaturated chondroitin and hyaluronan disaccharides with a 1,3-glycosidic bond. A single copy of UGL is included for each *Streptococcus* in the CAZy database, whereas *Flavobacterium johnsoniae* UW101, a probable close relative of *P. heparinus*, has two mutually homologous genes that code for UGL in the genome. Both the broad activity of streptococcal UGL toward various GAGs and its single gene copy in the genome provide evidence that the enzyme is unique in streptococci responsible for degrading uronate (ΔGlcUA and ΔIdoUA)-containing GAG disaccharides.

In addition to Arg-236, three other residues, Ser-365, Ser-368, and Lys-370, were shown to be involved in binding to sulfate groups in unsaturated chondroitin disaccharides, although the mutant (S368G) exhibited sufficient enzyme activity for binding. Therefore, the arrangement of four residues is identified to form the motif "R-//-S*XX*(S)*X*K" that is crucial for degradation of sulfated GAGs. The hyphen and slash marks in the motif indicate the presence of over 100 residues in the enzyme. The parentheses indicate that Ser-368 makes little contribution to enzyme activity. This motif is completely conserved in various UGLs from pathogenic species, including *Clostridium perfringens*, *Enterococcus faecalis*, *Erysipelothrix rhusiopathiae*, *Mycoplasma fermentas*, and streptococci such as *S. agalactiae*, *Streptococcus equi*, *S. pneumoniae*, *S. pyogenes*, and *Streptococcus suis*. In contrast, non-pathogenic soil isolates *Bacillus* sp. GL1 and *P. heparinus* include no such motif (Fig. 2-5). Some bacteria such as *P. heparinus* (21,22) and *Bacteroides thetaiotaomicron* (23) can assimilate GAGs as a sole carbon source. GAG lyases and UGL are also prerequisite for cell growth in these bacteria. Although the motif for degradation of sulfated GAGs is not induced in *P. heparinus* UGL, this bacterium produces some sulfatases involved in desulfation of GAGs (24,25).

The adhesion of pathogenic bacteria to mammalian cells is regarded as a primary mechanism of bacterial infection and plays an important role in the various secondary effects of the infectious process. GAGs present as an important component of the cell surface matrix are typical targets for microbial pathogens that invade host cells (26). Many specific interactions between pathogens and GAGs have been described previously (27). Among these bacteria, pathogenic enterococci and streptococci have been shown to interact with various human GAGs as follows. *E. faecalis* binds to heparin and heparan sulfate (28). A cell surface protein, alpha C protein, of *S. agalactiae* recognizes GAGs, likely heparin and heparan sulfate (29). *S. pneumoniae* binds to chondroitin sulfate, heparin, and heparan sulfate (30). *S. pyogenes* binds to dermatan sulfate, heparin, and heparan sulfate (31). These targeted GAGs (*i.e.* chondroitin sulfate, heparin, and heparan sulfate) are frequently sulfated;

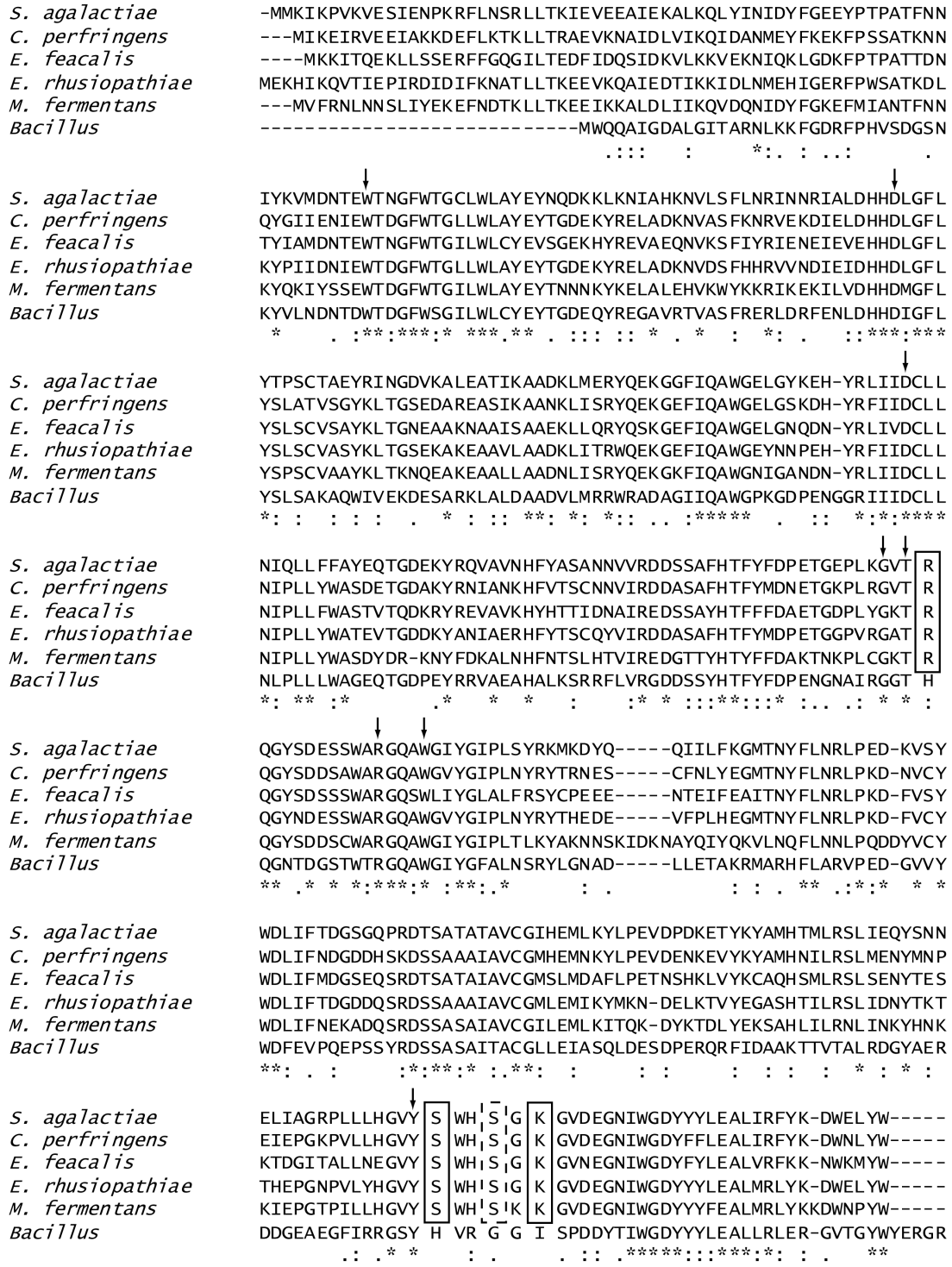

Figure 2-5. Multiple sequence alignment of bacterial UGLs. Bacteria with UGL genes were as follows: *S. agalactiae*, *Streptococcus agalactiae* NEM 316; *C. perfringens*, *Clostridium perfringens* ATCC 13124; *E. faecalis*, *Enterococcus faecalis* V583; *E. rhusiopathiae*, *Erysipelothrix rhusiopathiae* ATCC 19414; *M. fermentans*, *Mycoplasma fermentans* PG18; and *Bacillus*, *Bacillus* sp. GL1. The motif "R-//-S*XX*(S)*X*K" responsible for recognition of the sulfate groups of CΔ6S and CΔ4S is *boxed*. Other residues binding to the substrate are indicated by *arrows*. Identical and similar amino acid residues among proteins are denoted by *asterisks* and *dots*, respectively.

the sulfation level per the repeating disaccharide unit was found to be 2.4 for heparin, 1.0 for chondroitin sulfate, and 0.8–1.4 for heparan sulfate (27). Therefore, sulfate groups in GAGs are important for interactions between pathogenic bacteria and human cells. These interactions play a crucial role in promoting bacterial adhesion and invasion into human host cells. Although a large number of bacteria encode genes for UGL as well as for GAG lyases classified to the PL-8 in the CAZy database, it seems that pathogenic bacteria with the motif "R-//-S*XX*(S)*X*K" in the UGL sequence interact with sulfated GAGs. Little knowledge has been accumulated on the molecular mechanism for the interaction between streptococci and GAGs with the exception of the alpha C protein; therefore, further studies are necessary for elucidation of the involvement of UGL in streptococcal adhesion to mammal cells through degradation of GAGs.

In conclusion, this CHAPTER clarified the substrate recognition mechanism of streptococcal UGL through determination of crystal structure of the enzyme–substrate complex. This bacterial UGL can act on unsaturated heparin disaccharides, in addition to unsaturated chondroitin and hyaluronan disaccharides, because the enzyme triggers hydration of the vinyl ether group in unsaturated uronate residues (ΔGlcUA and ΔIdoUA). Examination of the active site structure of the sulfated substrate-bound UGL mutant and subsequent site-directed mutagenesis suggest that the motif "R-//-S*XX*(S)*X*K" in UGL is prerequisite for degradation of sulfated substrate.

References

1. Mori, S., Akao, S., Nankai, H., Hashimoto, W., Mikami, B., and Murata, K. (2003) A novel member of glycoside hydrolase family 88: overexpression, purification, and characterization of unsaturated β-glucuronyl hydrolase of *Bacillus* sp. GL1. *Protein Expr. Purif.* **29**, 77-84
2. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
3. Laemmli, U. K. (1970) Cleavage of structural proteins during assembly of head of bacteriophage T4. *Nature* **227**, 680-685
4. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326
5. Vagin, A. A., and Isupov, M. N. (2001) Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1451-1456
6. Bailey, S. (1994) The *CCP4* suite - programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50**, 760-763
7. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* **53**, 240-255
8. Emsley, P., and Cowtan, K. (2004) *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126-2132
9. Kabsch, W. (1976) Solution for best rotation to relate two sets of vectors. *Acta Crystallogr.*

Sect. A **32**, 922-923
10. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *PROCHECK* - a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291
11. DeLano, W. L. (2004) *The PyMOL Molecular Graphics System*. DeLano Scientific LLC, San Carlos, CA
12. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98**, 10037-10041
13. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The protein data bank. *Nucleic Acids Res.* **28**, 235-242
14. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
15. Hashimoto, W., Kobayashi, E., Nankai, H., Sato, N., Miya, T., Kawai, S., and Murata, K. (1999) Unsaturated glucuronyl hydrolase of *Bacillus* sp. GL1: novel enzyme prerequisite for metabolism of unsaturated oligosaccharides produced by polysaccharide lyases. *Arch. Biochem. Biophys.* **368**, 367-374
16. Gandhi, N. S., and Mancera, R. L. (2008) The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* **72**, 455-482
17. Itoh, T., Hashimoto, W., Mikami, B., and Murata, K. (2006) Crystal structure of unsaturated glucuronyl hydrolase complexed with substrate. *J. Biol. Chem.* **281**, 29807-29816
18. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557-559
19. Myette, J. R., Shriver, Z., Kiziltepe, T., McLean, M. W., Venkataraman, G., and Sasisekharan, R. (2002) Molecular cloning of the heparin/heparan sulfate Δ4,5 unsaturated glycuronidase from *Flavobacterium heparinum*, its recombinant expression in *Escherichia coli*, and biochemical determination of its unique substrate specificity. *Biochemistry* **41**, 7424-7434
20. Gu, K. N., Linhardt, R. J., Laliberte, M., Gu, K. F., and Zimmermann, J. (1995) Purification, characterization and specificity of chondroitin lyases and glycuronidase from *Flavobacterium heparinum*. *Biochem. J.* **312**, 569-577
21. Nader, H. B., Porcionatto, M. A., Tersariol, I. L. S., Pinhal, M. A. S., Oliveira, F. W., Moraes, C. T., and Dietrich, C. P. (1990) Purification and substrate specificity of heparitinase I and heparitinase II from *Flavobacterium heparinum*. Analyses of the heparin and heparan sulfate degradation products by ^{13}C NMR spectroscopy. *J. Biol. Chem.* **265**, 16807-16813
22. Tkalec, A. L., Fink, D., Blain, F., Zhang-Sun, G. Y., Laliberte, M., Bennett, D. C., Gu, K. F., Zimmermann, J. J. F., and Su, H. S. (2000) Isolation and expression in *Escherichia coli* of *cslA* and *cslB*, genes coding for the chondroitin sulfate degrading enzymes chondroitinase AC and chondroitinase B, respectively, from *Flavobacterium heparinum*. *Appl. Environ. Microbiol.* **66**, 29-35
23. Cheng, Q. O., Yu, M. C., Reeves, A. R., and Salyers, A. A. (1995) Identification and characterization of a *Bacteroides* gene, *csuF*, which encodes an outer membrane protein that is essential for growth on chondroitin sulfate. *J. Bacteriol.* **177**, 3721-3727
24. Myette, J. R., Soundararajan, V., Shriver, Z., Raman, R., and Sasisekharan, R. (2009) Heparin/heparan sulfate 6-*O*-sulfatase from *Flavobacterium heparinum* integrated structural and biochemical investigation of enzyme active site and substrate specificity. *J. Biol. Chem.* **284**, 35177-35188

25. Myette, J. R., Soundararajan, V., Behr, J., Shriver, Z., Raman, R., and Sasisekharan, R. (2009) Heparin/heparan sulfate *N*-sulfamidase from *Flavobacterium heparinum* structural and biochemical investigation of catalytic nitrogen sulfur bond cleavage. *J. Biol. Chem.* **284**, 35189-35200
26. Sawitzky, D. (1996) Protein glycosaminoglycan interactions: infectiological aspects. *Med. Microbiol. Immunol.* **184**, 155-161
27. Rostand, K. S., and Esko, J. D. (1997) Microbial adherence to and invasion through proteoglycans. *Infect. Immun.* **65**, 1-8
28. Sava, I. G., Zhang, F. M., Toma, I., Theilacker, C., Li, B. Z., Baumert, T. F., Holst, O., Linhardt, R. J., and Huebner, J. (2009) Novel interactions of glycosaminoglycans and bacterial glycolipids mediate binding of enterococci to human cells. *J. Biol. Chem.* **284**, 18194-18201
29. Aupérin, T. C., Bolduc, G. R., Baron, M. J., Heroux, A., Filman, D. J., Madoff, L. C., and Hogle, J. M. (2005) Crystal structure of the N-terminal domain of the group B *Streptococcus* alpha C protein. *J. Biol. Chem.* **280**, 18245-18252
30. Tonnaer, E. L. G. M., Hafmans, T. G., Van Kuppevelt, T. H., Sanders, E. A. M., Verweij, P. E., and Curfs, J. H. A. J. (2006) Involvement of glycosaminoglycans in the attachment of pneumococci to nasopharyngeal epithelial cells. *Microb. Infect.* **8**, 316-322
31. Frick, I. M., Schmidtchen, A., and Sjobring, U. (2003) Interactions between M proteins of *Streptococcus pyogenes* and glycosaminoglycans promote bacterial adhesion to host cells. *Eur. J. Biochem.* **270**, 2303-2311

CHAPTER III

Complete degradation of heparin and heparan sulfate by bacteria

SECTION 1

Structure–function relationship of heparan sulfate lyase

Distinct from chondroitin sulfate and hyaluronan with 1,3-glycosidic bonds between uronic acid and amino sugar residues, heparin and heparan sulfate contain 1,4-glycosidic bonds as described in INTRODUCTION. The differences between heparin and heparan sulfate are in their IdoUA and sulfate group contents, with heparin having more IdoUA and sulfate groups. *P. heparinus* isolated from soil has been extensively investigated as a heparin-degrading bacterium (1,2). This bacterium produces three heparin lyases as well as two chondroitin lyases, which are different in terms of their sequence similarity and substrate specificity (3,4). Family PL-13 heparin lyase I (HepA) prefers heparin with high degrees of sulfation, whereas family PL-12 heparin lyase III (HepC) is specific for heparan sulfate which includes sulfate group-free GlcUA-GlcNAc as a major disaccharide-repeating unit (Fig. 3-1-1a) (5). Thus, HepC is referred to as a heparan sulfate lyase. On the other hand, family PL-21 heparin lyase II (HepB) is active on both heparin and heparan sulfate.

Based on their biochemical and biophysical significance, *Pedobacter* HepA, HepB, and HepC have been thoroughly investigated by enzymology and site-directed mutagenesis (6-12). For example, each of 13 histidine residues of *Pedobacter* HepC (His-36, His-105, His-110, His-139, His-152, His-225, His-234, His-241, His-295, His-424, His-469, His-510, and His-539) has been substituted with Ala residue, and subsequently His-295 and His-510 have been demonstrated to be crucial for the HepC activity (13). The structure–function relationships of family PL-13 *B. thetaiotaomicron* HepA and family PL-21 *Pedobacter* HepB have also been studied by X-ray crystallography (14,15). No crystal structure of family PL-12 heparan sulfate lyase has been reported, although a heparinase III from *B. thetaiotaomicron* (*Bacteroides* HepC) has recently been analyzed structurally (16). Sequence identity between *Pedobacter* and *Bacteroides* HepCs is only 28%, although both are categorized into family PL-12. A significant difference in isoelectric point between both is observed (*Pedobacter* HepC, theoretical pI 9.0; *Bacteroides* HepC, theoretical pI

4.9) (Fig. 3-1-1b). There are no experimental data for structure-based functional analysis of amino acid residues that are crucial for HepC enzymatic activity.

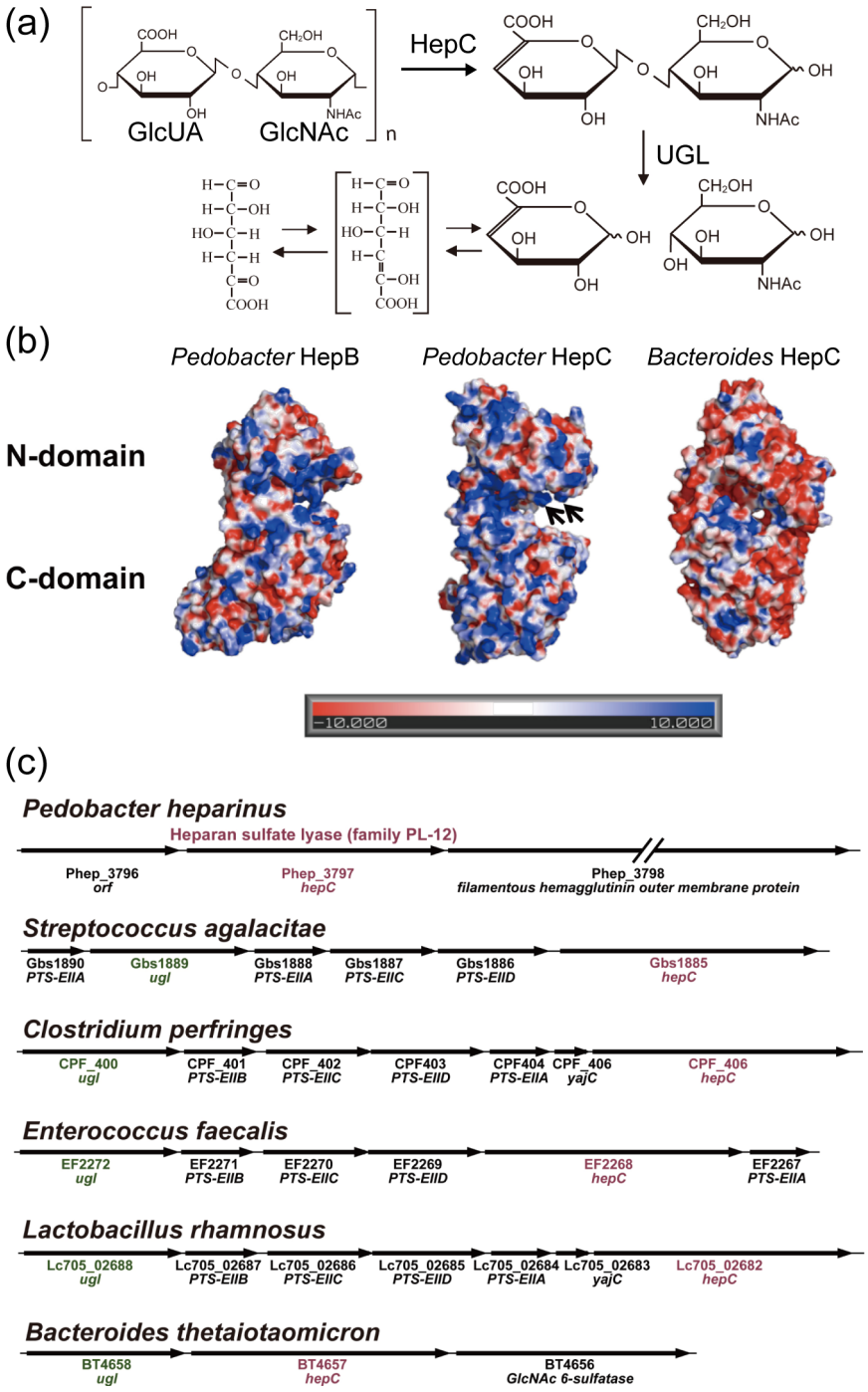

Figure 3-1-1. Heparin lyases. (a) Scheme for heparan sulfate degradation by HepC. Heparan sulfate includes sulfate group-free GlcUA-GlcNAc as a major disaccharide-repeating unit. (b) Electrostatic features on molecular surface model. *Left*, *Pedobacter* HepB; *center*, *Pedobacter* HepC; and *right*, *Bacteroides* HepC. Positive and negative charges are colored *blue* and *red*, respectively. Positively charged clusters at the active cleft of *Pedobacter* HepC are indicated by *arrows*. (c) Bacterial UGL gene clusters. PTS-EIIA–D: subunit domains of a phosphotransferase system for amino sugar import.

CHAPTERs I and II, together with previous analysis demonstrated that UGL was crucial for degrading unsaturated disaccharides generated from GAGs (Fig. 3-1-1a) (17). As described in CHAPTER I, the genetic cluster for UGL inducibly expressed in the presence of hyaluronan is in streptococcal genome. The cluster includes phosphotransferase system for incorporating the amino sugar GalNAc (Fig. 3-1-1c) (18,19). UGL gene cluster expression is important for degrading GAGs (hyaluronan and chondroitin) with 1,3-glycosidic bonds and for causing streptococcal infectious diseases (20). However, the involvement of the UGL gene cluster in degrading GAGs with 1,4-glycosidic bonds remains to be established.

In this CHAPTER, degradation of GAGs containing 1,4-glycosidic bonds by bacterial HepC and UGL was analyzed. SECTION 1 deals with the structure–function relationship of *Pedobacter* HepC by X-ray crystallography, site-directed mutagenesis, and differential scanning fluorimetry.

Materials and Methods

Molecular Cloning. An expression system for the signal peptide-truncated *Pedobacter* HepC (ORF ID in the bacterial genome database, Phep_3797) from *P. heparinus* NBRC 12017 (DSM 2366) purchased from the NBRC collection was constructed in *E. coli* cells as follows. To introduce the *Pedobacter* HepC gene into a pET21b expression vector, PCR was run using KOD-Plus polymerase (Toyobo), the genomic DNA of *P. heparinus* as a template, and two synthetic oligonucleotides as primers. The sequences of oligonucleotides with *Nde*I and *Xho*I sites added to their 5' regions for *Pedobacter* HepC are shown in Table 3-1-1. The pET21b vector was designed to express proteins with a hexahistidine (His_6)-tagged sequence at the C-terminus. The amplified truncated gene fragment was digested with *Nde*I and *Xho*I, and then ligated with *Nde*I and *Xho*I-digested pET21b. The resulting plasmid containing the truncated gene was designated pET21b-Phep_3797. In addition to the wild-type gene for *Pedobacter* HepC, a mutant gene for the double mutant (I29V/L657S) of *Pedobacter* HepC was unexpectedly amplified through the PCR procedure. Likewise, pET21b-*Lactobacillus rhamnosus* HepC homolog (*Lactobacillus* HepC) was constructed using *L. rhamnosus* NBRC 3425 purchased from the NBRC collection and two synthetic oligonucleotides (Table 3-1-1).

Protein Expression and Purification. E. coli strain BL21(DE3) or B834(DE3) (Novagen) was used as a host for *Pedobacter* HepC expression. For expression in *E. coli*, cells were aerobically precultured at 30°C in LB medium (21) supplemented with sodium ampicillin (0.1 mg ml^{-1}). For expression of *Pedobacter* HepC derivative with selenomethionine, *E. coli* cells were aerobically

cultured in a minimal medium (22) supplemented with 25 μg ml^{-1} of selenomethionine. When culture turbidity was approximately 0.5 at 600 nm, IPTG was added to the culture (0.1 mM) and the cells were cultured at 16°C for an additional 44 h. Unless otherwise specified, all purification procedures were done at 0–4°C. Recombinant *Pedobacter* HepC was purified from *E. coli* cells harboring pET21b-Phep_3797 to homogeneity through cell disruption by sonication followed using column chromatography with three different separation media, affinity (TALON, Clontech), cation-exchange (TOYOPEARL CM-650M), and gel filtration (Sephacryl S-200HR). Protein purity was confirmed using SDS-PAGE (23). Protein contents were determined by measuring the absorbance at 280 nm using a 1 cm path length cuvette and assuming that ε_{280} = 1.75 (*Pedobacter* HepC) corresponded to 1 mg ml^{-1}. Likewise, *Lactobacillus* HepC was expressed in *E. coli* strain BL21(DE3) cells and then purified to almost homogeneity.

Table 3-1-1. Primers

Primers for molecular cloning of HepC		
Pedobacter	forward	5′-GGCATATGCAAAGCTCTTCCATTACCAGGAAAGATTTT-3′
	reverse	5′-GGCTCGAGAGGAACCAACACAAGCTGTTGTTTTCCGTT-3′
Lactobacillus	forward	5′-GGCATATGGCGAAAATGACTTGGACCACGTACATTCAA-3′
	reverse	5′-GGCTCGAGTAACAATTGAAAATGTTCGATAACTTTGTG-3′
Primers for site-directed mutagenesis		
E237A	sense	5′-TATTTATCTACGCATTATGCCGCACAGGGAAACCACCGTTTATTT-3′
	antisense	5′-AAATAAACGGTGGTTTCCCTGTGCGGCATAATGCGTAGATAAATA-3′
Q238A	sense	5′-TTATCTACGCATTATGCCGAAGCGGGAAACCACCGTTTATTTGAA-3′
	antisense	5′-TTCAAATAAACGGTGGTTTCCCGCTTCGGCATAATGCGTAGATAA-3′
N240A	sense	5′-ACGCATTATGCCGAACAGGGAGCCCACCGTTTATTTGAAGCCCAA-3′
	antisense	5′-TTGGGCTTCAAATAAACGGTGGGCTCCCTGTTCGGCATAATGCGT-3′
H241A	sense	5′-CATTATGCCGAACAGGGAAACGCCCGTTTATTTGAAGCCCAACGC-3′
	antisense	5′-GCGTTGGGCTTCAAATAAACGGGCGTTTCCCTGTTCGGCATAATG-3′
Y294F	sense	5′-CAGTTTGAACTTTCACCAATTTTCCATGTAGCTGCCATCGATATC-3′
	antisense	5′-GATATCGATGGCAGCTACATGGAAAATTGGTGAAAGTTCAAACTG-3′
W350A	sense	5′-ACCCCTATGTTTGGAGATTCAGCGATTACAGATAAAAATTTCAGG-3′
	antisense	5′-CCTGAAATTTTTATCTGTAATCGCTGAATCTCCAAACATAGGGGT-3′
F423A	sense	5′-AAAGCCAGTCCTCCCGGAGAAGCTCATGCCCAGCCGGATAACGGG-3′
	antisense	5′-CCCGTTATCCGGCTGGGCATGAGCTTCTCCGGGAGGACTGGCTTT-3′
H424A	sense	5′-GCCAGTCCTCCCGGAGAATTTGCTGCCCAGCCGGATAACGGGACT-3′
	antisense	5′-AGTCCCGTTATCCGGCTGGGCAGCAAATTCTCCGGGAGGACTGGC-3′
Y450F	sense	5′-CCAGACGCCGGGGTATTTGTGTTTAGCGGCGACGAAGCCATCATG-3′
	antisense	5′-CATGATGGCTTCGTCGCCGCTAAACACAAATACCCCGGCGTCTGG-3′
Y590F	sense	5′-GAAGGAAAGGTATCTTATGTTTTCAATAAGGAGCTGAAAAGACCT-3′
	antisense	5′-AGGTCTTTTCAGCTCCTTATTGAAAACATAAGATACCTTTCCTTC-3′

Underlines in primers for molecular cloning and site-directed mutagenesis indicate the restriction and mutation sites, respectively.

Enzyme Assay. The enzyme was incubated in a 0.5 ml reaction mixture containing 0.01% substrate [heparan sulfate (Iduron), sodium heparin, chondroitin A, chondroitin C (Nacalai Tesque), or hyaluronan (Fluka)] and 50 mM Tris–HCl (pH 7.5). Activity was determined by monitoring the increase in absorbance at 235 nm. One unit of enzyme activity was defined as the amount of enzyme required to produce an increase of 1.0 in absorbance at 235 nm per minute.

Crystallization and X-ray Diffraction. For crystallization, purified native and selenomethionine derivative forms of *Pedobacter* HepC were concentrated to about 5 mg ml^{-1} by ultrafiltration with a Centriprep. *Pedobacter* HepC was crystallized at 20°C using the sitting-drop vapor-diffusion method. A mixture (50 μl) of 20% polyethylene glycol 3,000, 0.2 M calcium acetate, and 0.1 M Tris–HCl (pH 7.0) was used as a reservoir solution, and *Pedobacter* HepC (1 μl) was mixed with this reservoir solution (1 μl) to form a drop. Another reservoir solution [20% polyethylene glycol 3,350, 0.2 M sodium acetate, and 0.1 M Bis-Tris propane (pH 7.5)] was also suitable for crystallization of *Pedobacter* HepC. For cryoprotection, a protein crystal was soaked in the reservoir solution containing 20% glycerol. A crystal was picked up from the soaking solution with a mounted nylon loop (Hampton Research) and placed directly into a cold nitrogen gas stream at –173°C. For phasing, the derivative crystal with selenomethionine was prepared in a manner similar to that for the native crystal. X-ray diffraction images were acquired for the native and derivative crystals at –173°C under a nitrogen gas stream with a Quantum 315 CCD detector (ADSC) and synchrotron radiation (λ = 1.0000 Å for the native crystal or 0.9790 Å for the derivative crystal) at the BL38B1 station of SPring-8. Diffraction data were processed using the *HKL2000* program package (24). Data collection statistics for native and derivative crystals are summarized in Table 3-1-2.

Structure Determination and Refinement. The crystal structure of *Pedobacter* HepC was resolved using single-wavelength anomalous diffraction with the selenomethionine derivative crystal. Phase determination and initial model building for the *Pedobacter* HepC derivative were done using the *PHENIX* program (25). The *Coot* program (26) was used to manually modify the initial model. Initial rigid body refinement and several rounds of restrained refinement against the dataset were done using the *Refmac5* program (27). Water molecules were incorporated where the difference in density exceeded 3.0 σ above the mean and the $2F_o–F_c$ map showed a density of > 1.2 σ. Refinement continued until convergence at 2.20Å resolution. Final model quality was checked with *RAMPAGE* (28). Figures for ribbon plots were prepared using the *PyMol* program (29). Coordinates were obtained from the PDB (30). The atomic coordinates and structure factors of

Pedobacter HepC (ID: 4MMH) and its I29V/L657S mutant (ID: 4MMI) have been deposited in the RCSB PDB (http://www.rcsb.org/).

Site-directed Mutagenesis. To substitute Glu-237, Gln-238, Asn-240, His-241, Tyr-294, Trp-350, Phe-423, His-424, Tyr-450, and Tyr-590 of *Pedobacter* HepC with Ala, Ala, Ala, Ala, Phe, Ala, Ala, Ala, Phe, and Phe, respectively, 20 oligonucleotides were synthesized as described in Table 3-1-1. Site-directed mutagenesis was performed using plasmid pET21b-Phep_3797 as a template and synthetic oligonucleotides as sense and antisense primers using the methods described in a QuikChange site-directed mutagenesis kit, except that KOD-Plus polymerase was used for PCR. The resulting plasmids with mutations were used for mutant expression. Mutations were confirmed using DNA sequencing (31). *E. coli* host strain cells were transformed with each plasmid.

Differential Scanning Fluorimetry. Interactions between *Pedobacter* HepC and various saccharides were analyzed using differential scanning fluorimetry as described by Niesen *et al* (32), using a MyiQ2 real-time PCR instrument (Bio-Rad). Heparin di- and tetrasaccharides (Iduron), ΔGellan (33), and GlcNAc were used as ligand sacharides. Fluorescence derived from a commercial dye, SYPRO Orange (Invitrogen), was monitored using filters provided with the PCR instrument (excitation at 492 nm and emission at 610 nm). A reaction mixture (20 μl) that included *Pedobacter* HepC (0.232 mg ml^{-1} = 3.1 μM), each saccharide (0–1.0 mM), SYPRO Orange (1,000-fold dilution), and Tris–HCl (20 mM, pH 7.5) was subjected to heat treatment. The temperature was increased from 25°C to 95°C by 0.5°C $cycle^{-1}$ (10 sec $cycle^{-1}$) for a total of 141 cycles. The fluorescence emitted from SYPRO Orange after it bound to a denatured protein was measured. The fluorescence profile was obtained by plotting the relative fluorescence unit (RFU) at each temperature. The resulting fluorescence profile was analyzed using the *iQ5* program (Bio-Rad) and the midpoint of the increase in the profile was defined as the melting temperature.

Results and Discussion

Crystallization and Structure Determination

Pedobacter HepC (ORF ID, Phep_3797) consists of 659 residues with a signal peptide of 24 residues (Met-1–Ala-24). The recombinant *Pedobacter* HepC protein expressed in *E. coli* cells did not include the signal peptide (Thr-2–Ala-24), but did include eight additional residues (Leu–Glu–His–His–His–His–His–His) at its C-terminus. This indicated that the recombinant enzyme contained 644 residues with an N-terminal sequence of Met-1–Gln-25–Ser-26–Ser-27 and a C-terminal sequence of Leu-657–Val-658–Pro-659–Leu-660–Glu-661–His-662–His-663–His-664–

His-665–His-666–His-667. The recombinant *E. coli* cells produced significant amounts of the enzyme in a soluble form and purified *Pedobacter* HepC (19.4 mg) was obtained from 4.5 litter of LB culture (Fig. 3-1-2a). Similar to the native enzyme from *P. heparinus* (3), the recombinant enzyme was specific for heparan sulfate; specific activities of 23 ± 3 U mg^{-1} for heparan sulfate and 1.1 ± 0.2 U mg^{-1} for sodium heparin. Thus, to analyze its three-dimensional structure, the recombinant enzyme was subjected to crystallization.

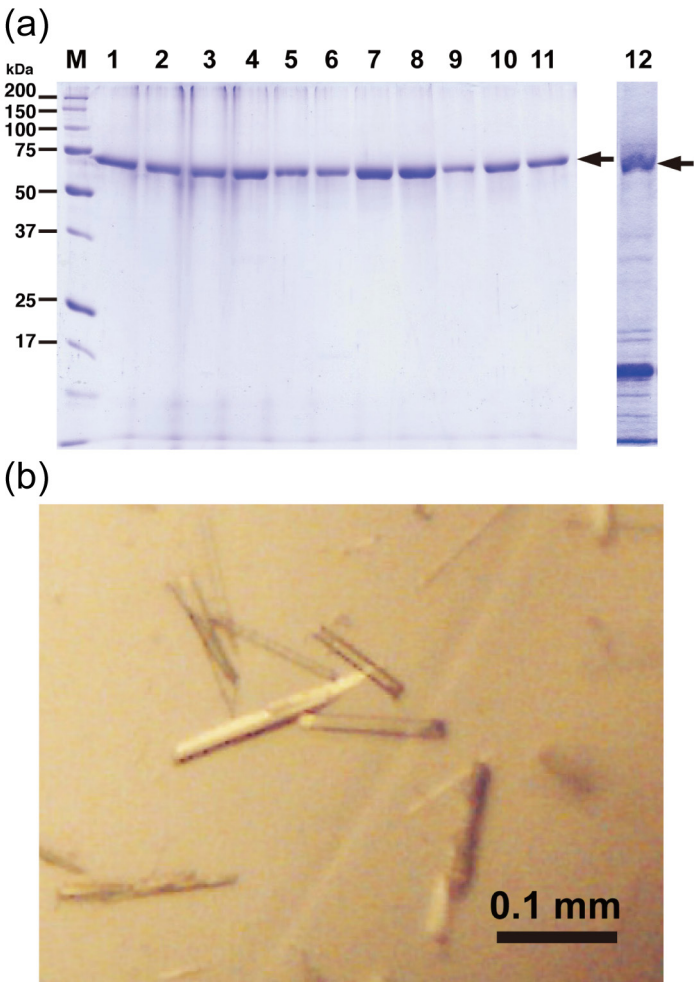

Figure 3-1-2. Purification and crystallization of bacterial HepC. (a) SDS-PAGE profile. *Lane M*, molecular weight standards (from top): synthetic polypeptides with molecular weights of 200,000, 150,000, 100,000, 75,000, 50,000, 37,000, 25,000, and 17,000. *Lane 1*, *Pedobacter* HepC wild-type; *lane 2*, *Pedobacter* HepC E237A; *lane 3*, *Pedobacter* HepC Q238A; *lane 4*, *Pedobacter* HepC N240A; *lane 5*, *Pedobacter* HepC H241A; *lane 6*, *Pedobacter* HepC Y294F, *lane 7*, *Pedobacter* HepC W350A; *lane 8*, *Pedobacter* HepC F423A; *lane 9*, *Pedobacter* HepC H424A; *lane 10*, *Pedobacter* HepC Y450F; *lane 11*, *Pedobacter* HepC Y590F; and *lane 12*, *Lactobacillus* HepC. (b) *Pedobacter* HepC crystals in droplet A.

Two different stick-shaped crystals of *Pedobacter* HepC were found under different crystallization conditions. One was in droplet A comprising 20% polyethylene glycol 3,000, 0.2 M calcium acetate, and 0.1 M Tris–HCl (pH 7.0) (Fig. 3-1-2b), and the other was in droplet B consisting of 20% polyethylene glycol 3,350, 0.2 M sodium acetate, and 0.1 M Bis-Tris propane (pH 7.5). Within three weeks, both crystals in these drops grew at 20°C to a size of > 0.1 mm. Diffraction images were acquired at up to 2.20 Å resolution for the crystal (wild-type *Pedobacter* HepC) in droplet A and 2.40 Å resolution for the crystal (the I29V/L657S mutant) in droplet B after cryoprotection. The results of X-ray data collection are listed in Table 3-1-2.

The refined model for *Pedobacter* HepC comprised 637 residues and 131 water molecules for a protein molecule in an asymmetric unit. All amino acid residues (Ile-29–//–Pro-659–Leu-660–Glu-661–His-662–His-663–His-664–His-665), except for the N-terminal residues (Met-1–Gln-25–Ser-26–Ser-27–Ser-28) and a portion of the C-terminal His-tagged sequence (His-666–His-667), could be assigned well in the $2F_o-F_c$ map.

The final overall *R*-factor for the refined model was 0.195 at up to 2.20 Å resolution. The final free *R*-factor calculated using 5% of randomly selected data was 0.244. The final r.m.s.d. from the standard geometry was 0.005 Å for bond lengths and 0.995° for bond angles. On the basis of the results of Ramachandran plot analysis for which the stereochemical correctness of the backbone structure was indicated by the φ and ψ torsion angles (34), most of the non-glycine residues were within the favored regions, and the other residues were in additional and generously allowed regions. The structure of the I29V/L657S mutant crystallized in droplet B was determined by molecular replacement using refined wild-type coordinates as an initial model, although the mutant was also crystallized in droplet A. Refinement statistics are listed in Table 3-1-2. Although compositions of droplets A and B mutually differ, properties of crystals formed in both droplets are identical. The structure of I29V/L657S was also essentially identical with that of the wild-type *Pedobacter* HepC other than the interdomain relationship described later.

Overall Structure

The overall structure (Fig. 3-1-3a) indicated that *Pedobacter* HepC was composed of two globular domains (N- and C-terminal domains) that formed α- and β-structures, respectively. The N-terminal domain comprised 350 residues from Ile-29 to Ala-378 and one metal ion and was predominantly composed of 15 α-helices (HA0–HA3, HA3', HA4, HA4', and HA5–HA12), twelve of which formed an α/α-barrel structure (Fig. 3-1-3b). The C-terminal domain comprised 268 residues from Lys-392 to Pro-659 and one metal ion and consisted of one helix (HB1) and 20

β-strands (SA1–SA4, SB1–SB4, SC1–SC9, and SD1–SD3) arranged in four antiparallel β-sheets (sheets A–D; Fig. 3-1-3c). A peptide linker comprising 13 residues from Thr-379 to Ser-391 connected the N- and C-terminal domains.

Table 3-1-2. Data collection and refinement statistics

	Pedobacter HepC wild-type	*Pedobacter* HepC with selenomethionine	I29V/L657S
Data collection			
Wavelength (Å)	1.0000	0.9790	1.0000
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å)	a = 41.2, b = 104.1, c = 143.3	a = 42.7, b = 105.4, c = 145.8	a = 44.7, b = 102.9, c = 149.2
Resolution limit (Å)	50.00–2.20 (2.28–2.20)[a]	50.00–2.70 (2.80–2.70)	50.00–2.40 (2.49–2.40)
No. of measured reflections	257,911	149,613	210,361
No. of unique reflections	32,238	18,682	27,715
Redundancy	8.0 (8.1)	8.0 (8.1)	7.6 (7.3)
Completeness (%)	99.9 (100)	100 (100)	99.7 (99.5)
I/σ (I)	34.5 (6.8)	26.7 (6.2)	38.3 (4.1)
R_{merge}	0.108 (0.424)	0.131 (0.496)	0.078 (0.398)
Refinement			
R_{work}	0.195		0.205
R_{free}	0.244		0.249
No. of molecules/asymmetric unit	1		1
No. of non-hydrogen atoms			
Protein	5,214		5,201
Calcium ions	2		2
Water molecules	131		61
Average isotropic B-factor (Å^2)			58.7
Protein	32.0		59.0
Calcium ions	23.5		52.2
Water molecules	27.3		47.9
r.m.s.d. from ideal			
Bond lengths (Å)	0.005		0.006
Bond angles (deg)	0.995		1.079
Ramachandran plot (%)			
Favored regions	98.0		98.1
Allowed regions	2.0		1.9
Outlier region	0		0

[a] Data on highest shells are in parenthesis.

Two metal ions were included in the N- and C-terminal domains (Fig. 3-1-3a). Because droplet A included calcium acetate, these metal ions were probably calcium ions. Although ethylenediaminetetraacetic acid (EDTA) barely inhibited the enzyme activity of *Pedobacter* HepC

in this assay, calcium ions are known to be required for its activity (7). Two calcium ion-binding sites (Leu-390–Gly-405 and Arg-576–Asn-591) were postulated for the *Pedobacter* HepC molecule based on its primary structure (7). However, these putative sites did not correspond to calcium-binding sites in the crystal structure. The four atoms (Nδ1 of His-241, Oε1 of Glu-245, and two oxygen atoms of water molecules) in the N-terminal domain were coordinated to a calcium ion. The distance between this calcium ion and the oxygen atoms ranged from 2.1 to 2.7 Å (average: 2.3 Å). On the other hand, six atoms (Oε1 of Gln-426, Oδ1 of Asp-444, Nε2 of His-469, and three oxygen atoms of water molecules) were coordinated to a calcium ion bound to the C-terminal domain, and the coordination geometry comprised a distorted octahedron. The distance between this calcium ion and the oxygen atoms ranged from 2.1 to 2.4 Å (average: 2.2 Å).

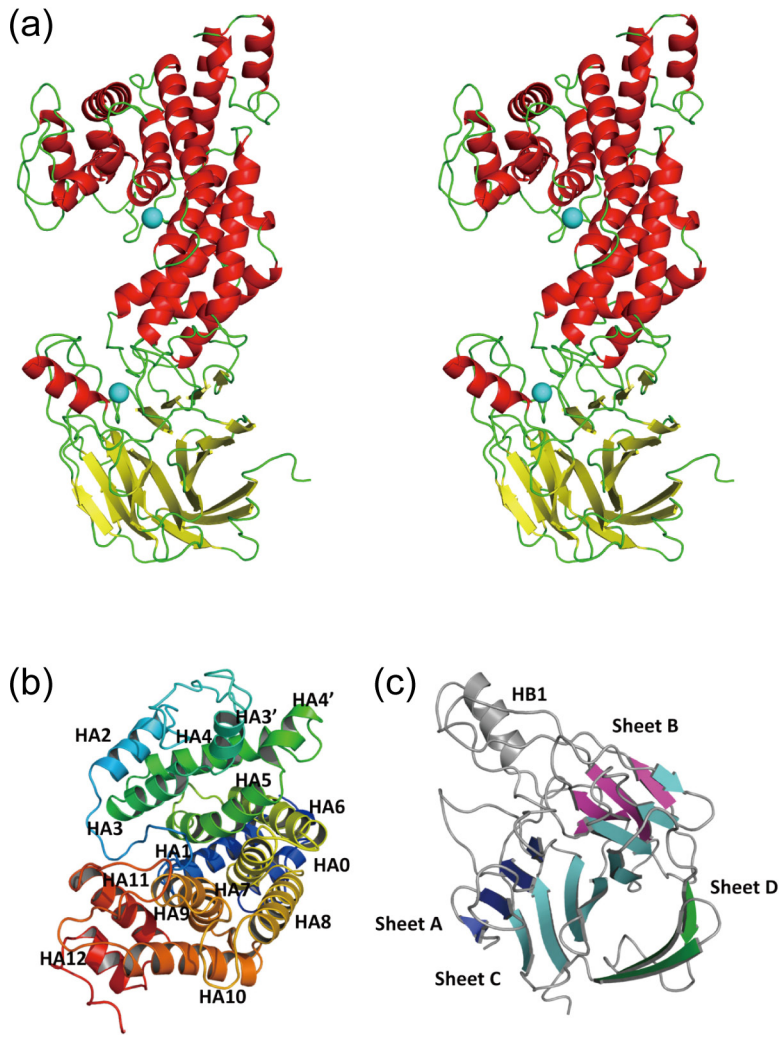

Figure 3-1-3. Structure of *Pedobacter* HepC. (a) Overall structure (stereo-diagram) (*ball* colored *cyan*: calcium ion). Colors denote secondary structure elements (*red*: α-helices; *yellow*: β-strands; *green*: loops and coils). (b) N-terminal α/α-barrel composed of twelve α-helices. (c) C-terminal antiparallel β-sheets (Sheets A–D).

Structural homologs of *Pedobacter* HepC were found in the PDB using the *DALI* program (35). Several proteins with N-terminal α/α-barrels and C-terminal antiparallel β-sheets as a basic scaffold were found to be structurally homologous with *Pedobacter* HepC (Table 3-1-3). The overall structure of *Pedobacter* HepC was most similar to the recently determined structure of family PL-12 HepC from *B. thetaiotaomicron* (PDB ID: 4FNV, Z = 50.3) (16) with an r.m.s.d. of 4.6 Å for 659 C^{α} atoms, although the level of sequence identity between these two was not very high (28%). *Pedobacter* HepC was also structurally similar to other polysaccharide lyases, such as family PL-15 alginate lyase Atu3025 (36) from *Agrobacterium tumefaciens* (PDB ID: 3A0O, Z = 25.0), family PL-21 *Pedobacter* HepB (15) (PDB ID: 3E80, Z = 24.2), and family PL-8 chondroitin AC lyase from *Arthrobacter aurescens* (PDB ID: 1RW9, Z = 21.8). A superimposition of the crystal structures of *Pedobacter* and *Bacteroides* HepCs is shown in Fig. 3-1-4a. The distance between the two domains of *Pedobacter* HepC was greater than that for *Bacteroides* HepC, which indicated that *Pedobacter* HepC exhibited an open form with two domains mutually separated.

Table 3-1-3. Structure-based homology with *Pedobacter* HepC

Z score	Description	r.m.s.d.	PDB ID
50.3	HepC from *B. thetaiotaomicron*	4.6	4FNV
25.0	Alginate lyase Atu3025 from *A. tumefaciens*	4.9	3A0O
24.2	HepB from *P. heparinus*	6.2	3E80
21.9	Chondroitin AC lyase from *Arthrobacter aurescens*	4.6	1RW9
20.8	Chondroitin AC lyase from *Streptomyces coelicolor*	4.5	2WDA
20.8	Chondroitin sulfate lyase ABC from *B. thetaiotaomicron*	4.6	2QLF
20.4	Chondroitinase AC lyase from *P. heparinus*	5.7	1CB8
20.3	Hyaluronate lyase from *S. agalactiae*	5.0	1LXM
19.0	Hyaluronate lyase from *S. pneumoniae*	5.2	1N7P
18.4	Xanthan lyase from *Bacillus* sp. GL1	4.2	1J0M

A large number of polysaccharide lyases with α/α-barrels were categorized into the superfamily of chondroitin AC/alginate lyase in "α/α toroid" fold in the SCOP database (http://scop.mrc-lmb.cam.ac.uk/scop/). *Pedobacter* HepC will be structurally classified into this "α/α toroid" fold in the SCOP database.

Active Site

The catalytic reactions of polysaccharide lyases can be divided into three steps as follows (37). (i) Positively charged residues (**Residue A**) stabilize or neutralize the negative charge on the C-6 carboxylate anion. (ii) A general base catalyst (**Residue B**) abstracts the proton from C-5 of the uronic acid residue. (iii) A general acid catalyst (**Residue C**) donates a proton to the glycosidic bond

to be cleaved. To assess the enzyme catalytic mechanism, enzyme crystals in complexes with heparin oligosaccharides were desired, although no crystal was obtained.

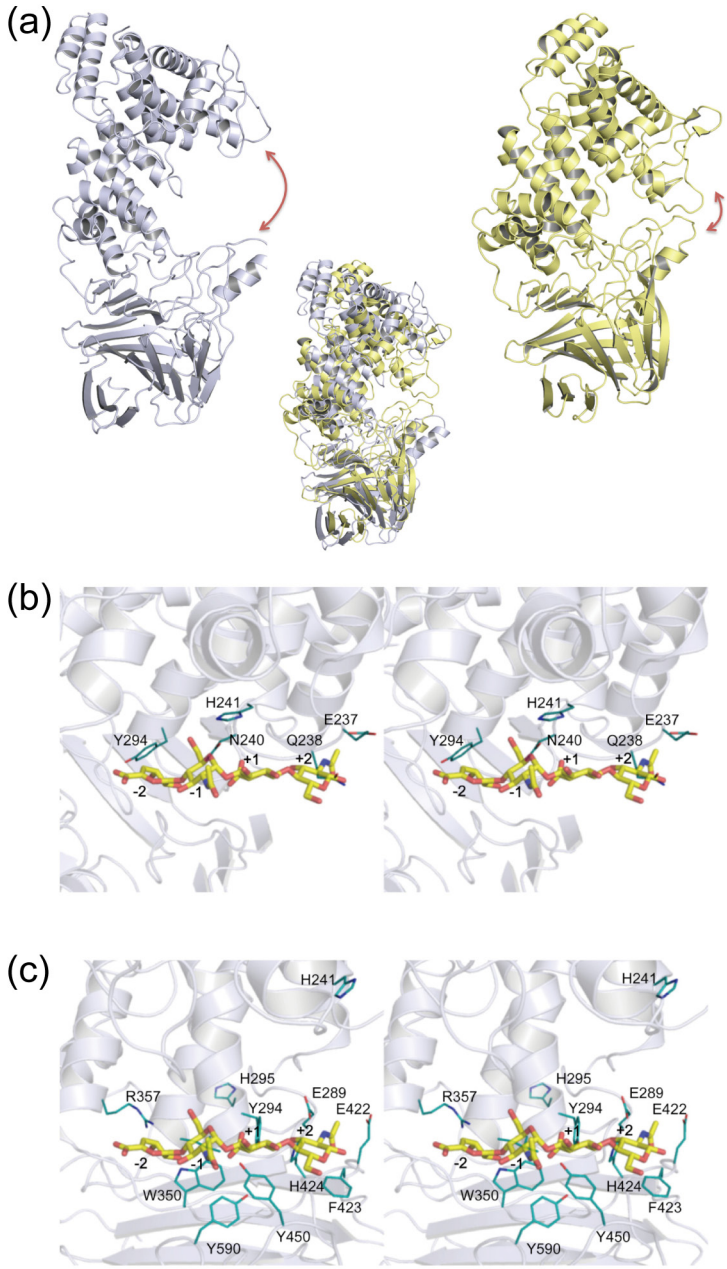

Figure 3-1-4. Structural comparisons. (a) Superimposition of *Pedobacter* HepC (*steel blue*) and *Bacteroides* HepC (*yellow*). (b) Superimposition of the N-terminal domains of *Pedobacter* HepC and *Pedobacter* HepB/tetrasaccharide (stereo-diagram). (c) Superimposition of the C-terminal domain of *Pedobacter* HepC and the central subdomain of *Pedobacter* HepB/tetrasaccharide (stereo-diagram). The tetrasaccharide (ΔGlcUA-GlcNAc-GlcUA-GlcNAc) from coordinates of *Pedobacter* HepB/tetrasaccharide (PDB ID, 3E7J) is colored *yellow* (carbon atom), *red* (oxygen atom), and *blue* (nitrogen atom). Probable residues of *Pedobacter* HepC crucial for binding tetrasaccharide are colored *cyan* (carbon atom), *red* (oxygen atom), and *blue* (nitrogen atom).

Because *Pedobacter* HepC was structurally similar to *Pedobacter* HepB complexed with heparin tetrasaccharide (PDB ID: 3E7J) (38), the active site in *Pedobacter* HepC was analyzed through superimposition of *Pedobacter* HepC on *Pedobacter* HepB/tetrasaccharide. In the case of *Pedobacter* HepB, the tetrasaccharide is bound to the cleft that is formed between the N-terminus and the central subdomains, which correspond to the N- and C-terminal domains of *Pedobacter* HepC, respectively (38). Because of the open form of *Pedobacter* HepC, two N- and C-terminal domains were analyzed separately. The N-terminal domain of *Pedobacter* HepC was superimposed on the N-terminal subdomain of *Pedobacter* HepB/tetrasaccharide (Fig. 3-1-4b), and the C-terminal domain of *Pedobacter* HepC was superimposed on the central subdomain of *Pedobacter* HepB/tetrasaccharide (Fig. 3-1-4c).

On the basis of the complex structure of *Pedobacter* HepB/tetrasaccharide, the following reaction mechanism of *Pedobacter* HepB catalysis was proposed (38). His-406 functions as a neutralizer (**Residue A**) for the substrate carboxyl group, and Tyr-257 successively functions as a general base and acid (**Residues B** and **C**) for GlcUA-containing heparin. His-202 is also considered to be crucial for enzyme catalysis as a general base (**Residue B**) for a substrate that contains IdoUA.

Structural comparisons indicated that His-202, Tyr-257, and His-406 of *Pedobacter* HepB corresponded to His-241, Tyr-294, and His-424 of *Pedobacter* HepC. In the case of *Pedobacter* HepC, His-241 appeared to be slightly farther from the tetrasaccharide (Fig. 3-1-4c). This was due to the open form of *Pedobacter* HepC, and His-241 was located near the tetrasaccharide in the possible closed conformation, as seen in the superimposition of both N-terminal domains of *Pedobacter* HepC and *Pedobacter* HepB/tetrasaccharide (Fig. 3-1-4b). Three *Pedobacter* HepC mutants, H241A, Y294F, and H424A, were constructed by site-directed mutagenesis, purified to homogeneity (Fig. 3-1-2a), and used for enzyme assays. The specific activities of H241A, Y294F, and H424A toward heparan sulfate are listed in Table 3-1-4. These mutants had significantly reduced enzyme activity compared with the wild-type enzyme. In particular, Tyr-294 and His-424 were suggested to be crucial for enzyme activity as with *Pedobacter* HepB. The mutation at His-241 has also been demonstrated to reduce the turnover number (k_{cat}) of the enzyme (13).

Other than putative catalytic residues (Tyr-294 and His-424), seven residues (Glu-237, Gln-238, Asn-240, Trp-350, Phe-423, Tyr-450, and Tyr-590) located at the active site (Fig. 3-1-4b, c) were substituted with an Ala or Phe residue by site-directed mutagenesis. Seven mutants (E237A, Q238A, N240A, W350A, F423A, Y450F, and Y590F) were purified to homogeneity (Fig. 3-1-2a)

and used for enzyme assays. The specific activities of these mutants toward heparan sulfate are also listed in Table 3-1-4. Among these mutants, N240A had significantly reduced enzyme activity, which suggested that Asn-240 in conjunction with Tyr-294 and His-424 was involved in enzyme activity as a stabilizer or catalyst (**Residue A**, **B**, or **C**). In fact, these Asn, Tyr, and His residues were conserved in all three heparan sulfate lyases from *P. heparinus*, *B. thetaiotaomicron*, and *Bacteroides stercoris* that have been well characterized (3,16,39). The importance of these residues has also been reported for *B. thetaiotaomicron* HepC (16).

Table 3-1-4. Specific activity of *Pedobacter* HepC wild-type and mutant enzymes toward heparan sulfate

Proteins	Specific activity (U mg^{-1})[a]	Relative activity (%)
Wild-type	23 ± 3	100
I29V/L657S	9.6 ± 1.5	43
E237A	9.2 ± 2.3	41
Q238A	20 ± 3	89
N240A	0.69 ± 0.21	3.1
H241A	6.7 ± 1.4	30
Y294F	0.64 ± 0.14	2.8
W350A	6.1 ± 1.1	27
F423A	9.6 ± 1.0	42
H424A	0.070 ± 0.018	0.31
Y450F	1.3 ± 0.1	5.6
Y590F	1.8 ± 0.4	7.7

[a] Each data represents the average of triplicate individual experiments (mean ± standard deviations).

Two histidine residues, His-295 and His-510, are suggested to be crucial for the enzyme activity of *Pedobacter* HepC by site-directed mutagenesis (13). His-295 constituted the active site in the crystal structure (Fig. 3-1-4c), although His-510 was situated far from the active cleft. The crystal structure of this enzyme in complex with heparin-derived saccharides will be needed to elucidate mechanisms of *Pedobacter* HepC for its catalytic reaction and substrate recognition.

Interactions between Pedobacter HepC and Heparin Oligosaccharides

To analyze its substrate binding, *Pedobacter* HepC was subjected to differential scanning fluorimetry based on changes in protein stability by ligand binding in the presence of a dye, SYPRO Orange. Various saccharides were used as the ligand at concentrations ranging from 0 to 1 mM. The fluorescence of the dye bound to denatured proteins was measured during heat treatment from 25 to 95°C. The fluorescence profile of *Pedobacter* HepC in the presence of heparin oligosaccharides, particularly tetrasaccharides, significantly shifted to a lower temperature as compared with that in

the absence of saccharides (Fig. 3-1-5). Melting temperatures, T_m, of *Pedobacter* HepC in the absence and presence of saccharides were determined as the midpoints of the increases in the fluorescence profiles (Table 3-1-5). Among various saccharides, heparin di- and tetrasaccharides showed effects by lowering the T_m. The T_m shifted to a lower temperature with an increased concentration of added heparin tetrasaccharide. Even a small amount of tetrasaccharide at a concentration of < 0.01 mM significantly reduced the T_m. This suggested that *Pedobacter* HepC had an affinity for heparin oligosaccharides and that ligand-bound *Pedobacter* HepC was thermally less stable than the ligand-free protein. The difficulty with preparing a ligand-bound enzyme crystal may have been due to the instability of the structure of this complex.

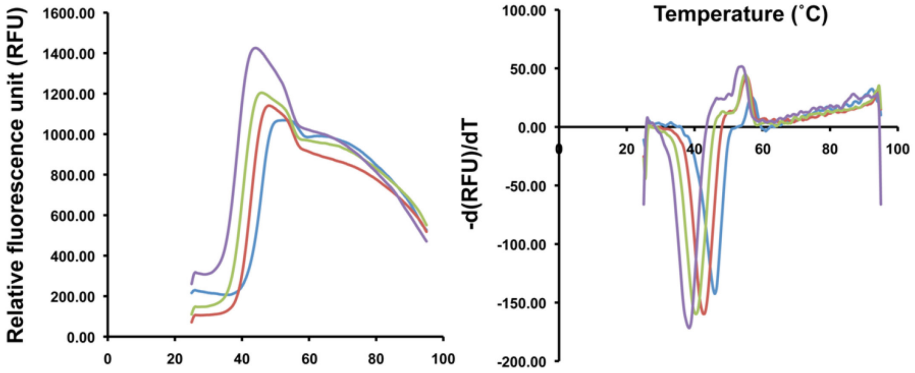

Figure 3-1-5. Substrate binding analysis by differential scanning fluorimetry. *Left panel*: fluorescence profile of *Pedobacter* HepC with heparin tetrasaccharide (*blue*: 0 mM; *brown*: 0.01 mM; *green*: 0.1 mM; *purple*: 1 mM). The tetrasaccharide purchased from Iduron is obtained mainly from the disaccharide units of IdoUA and GlcN with sulfate groups at C-2 and *N* positions, respectively, and the nonreducing uronate residue has a C4–C5 double bond. *Right panel*: negative derivative curve plot obtained from the fluorescence profile.

The shift in T_m suggested a conformational change in *Pedobacter* HepC through binding to heparin oligosaccharides. Because the mutants N240A and H241A exhibited lower enzyme activity, Asn-240 and His-241 were considered to be involved in substrate binding by conversion of the open form to the closed form (Fig. 3-1-4b, c). This conversion was supported by another crystal structure of the enzymatically active *Pedobacter* HepC double mutant (I29V/L657S) in the relatively closed form (Fig. 3-1-6). Although properties of the mutant crystal formed in droplet B were comparable with those of the wild-type enzyme (Table 3-1-2), the N- and C-terminal domains of the mutant were close to each other. On the basis of the results of differential scanning fluorimetry and site-directed mutagenesis, *Pedobacter* HepC is suggested to exhibit domain dynamics by substrate binding and product release.

Table 3-1-5. Temperature-denaturation transitions for *Pedobacter* HepC

Ligand	Concentration (mM)	T_m (°C)
None		45.44
Heparin tetrasaccharide	0.00001	45.44
	0.00005	45.01
	0.0001	44.79
	0.0005	43.93
	0.001	43.72
	0.005	42.64
	0.01	42.21
	0.05	42.21
	0.1	40.48
	0.5	40.06
	1	39.40
Heparin disaccharide	0.01	44.15
	0.1	43.93
	0.5	43.72
	1	43.50
ΔGellan	0.01	45.83
	0.1	45.19
	1	44.97
GlcNAc	0.01	44.53
	0.1	44.31
	1	44.31

As shown in Fig. 3-1-1b, positively charged clusters derived from basic residues (Lys-83, Lys-111, Arg-140, and Arg-188) are located at the substrate-binding cleft of *Pedobacter* HepC. All of them lie at the N-terminal side of the cleft, and the C-terminal side includes no basic residues, suggesting that acidic substrate of heparan sulfate first binds to *Pedobacter* HepC at the N-terminal side, followed by a conformational change to the closed form. The three positively charged residues, Lys-111, Arg-140, and Arg-188, of *Pedobacter* HepC correspond to Lys-101, Arg-131, and Arg-195, respectively, of *B. stercoris* HepC, although these basic residues are hardly conserved in *B. thetaiotaomicron* HepC and there is a significant difference in electrostatic features on overall molecular surface between *Pedobacter* and *Bacteroides* HepCs (Fig. 3-1-1b). Compared with the surface structure of *Pedobacter* HepC, *Pedobacter* HepB preferring sulfated substrate shows additional positive charge at the C-terminal side of the cleft. The difference in molecular surface charge may be a structural determinant for the substrate specificity of *Pedobacter* HepC.

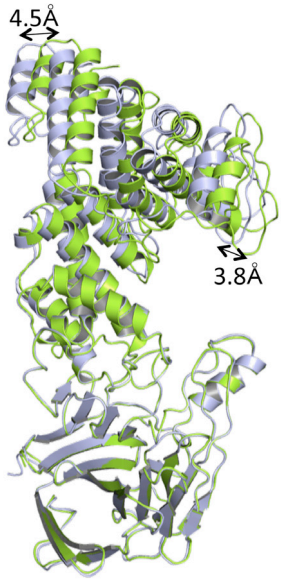

Figure 3-1-6. Domain dynamics. In the crystal structure of the *Pedobacter* HepC mutant (I29V/L657S) (*green*), the N- and C-terminal domains are close to each other as compared with those of *Pedobacter* HepC wild-type enzyme (*steel blue*). The structure of the mutant crystal formed in droplet B was determined at 2.40 Å resolution through molecular replacement by using the structure of *Pedobacter* HepC wild-type enzyme as a search model.

Gene Distribution

The streptococcal UGL gene cluster is involved in hyaluronan degradation (CHAPTER I) (40). Although the *Pedobacter* HepC (ID: Phep_3797) gene is located far from the genes that encode for 13 putative UGLs (IDs: Phep_0790, 0960, 1296, 2238, 2649, 2694, 2830, 2839, 2845, 3864, 3866, 3963, and 4097) in the genome of *P. heparinus* (41), a HepC homologous protein was found to be encoded downstream from the UGL gene cluster in normal flora and pathogenic bacteria, such as streptococci, clostridia, enterococci, lactobacillus, and bacteroides (Fig. 3-1-1c). CHAPTER II indicated that streptococcal UGL could degrade unsaturated disaccharides derived from GAGs with 1,4- as well as 1,3-glycosidic bonds. Analysis of streptococcal HepC homolog through its expression in *E. coli* cells was failed. Thus, the *Lactobacillus* HepC (42) encoded in the vicinity of UGL gene was expressed in *E. coli* cells and purified (Fig. 3-1-2a).

Using gel filtration column chromatography, *Lactobacillus* HepC with 641 amino acid residues expressed in a soluble form was a monomeric protein with a molecular mass of about 75 kDa. Similar to *Pedobacter* HepC, *Lactobacillus* HepC was active on heparin, and particularly specific for heparan sulfate, but was inert on hyaluronan and chondroitin with 1,3-glycosidic bonds. On the basis of all of the results described above, the organization of the UGL gene cluster is probably crucial for degrading GAGs with 1,4- as well as 1,3-glycosidic bonds.

The streptococcal UGL gene cluster was shown to be involved in the expression of bacterial diseases (CHAPTER I) (20). Because the architectures of mammalian extracellular matrices are composed of various complex GAGs, such as hyaluronan, chondroitin, heparin, heparan sulfate, etc., HepC encoded as a component of the UGL gene cluster in the bacterial genome is probably essential for degrading GAGs with 1,4-glycosidic bonds during the interactions between mammalian hosts and bacteria.

In conclusion, *Pedobacter* HepC comprised an N-terminal α/α-barrel domain and a C-terminal antiparallel β-sheet domain, both of which were flexible by hinge bending. Through its binding to a substrate, this enzyme seems to be converted to the closed form with access to its two domains. The active site in *Pedobacter* HepC is probably located in the deep cleft at the interface between these two domains.

References

1. Payza, A. N., and Korn, E. D. (1956) Bacterial degradation of heparin. *Nature* **177**, 88-89
2. Galliher, P. M., Cooney, C. L., Langer, R., and Linhardt, R. J. (1981) Heparinase production by *Flavobacterium heparinum*. *Appl. Environ. Microbiol.* **41**, 360-365
3. Lohse, D. L., and Linhardt, R. J. (1992) Purification and characterization of heparin lyases from *Flavobacterium heparinum*. *J. Biol. Chem.* **267**, 24347-24355
4. Pojasek, K., Shriver, Z., Kiley, P., Venkataraman, G., and Sasisekharan, R. (2001) Recombinant expression, purification, and kinetic characterization of chondroitinase AC and chondroitinase B from *Flavobacterium heparinum*. *Biochem. Biophys. Res. Commun.* **286**, 343-351
5. Gandhi, N. S., and Mancera, R. L. (2008) The structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* **72**, 455-482
6. Sasisekharan, R., Leckband, D., Godavarti, R., Venkataraman, G., Cooney, C. L., and Langer, R. (1995) Heparinase I from *Flavobacterium heparinum* - the role of the cysteine residue in catalysis as probed by chemical modification and site-directed mutagenesis. *Biochemistry* **34**, 14441-14448
7. Godavarti, R., and Sasisekharan, R. (1996) A comparative analysis of the primary sequences and characteristics of heparinases I, II, and III from *Flavobacterium heparinum*. *Biochem. Biophys. Res. Commun.* **229**, 770-777
8. Godavarti, R., and Sasisekharan, R. (1998) Heparinase I from *Flavobacterium heparinum* - role of positive charge in enzymatic activity. *J. Biol. Chem.* **273**, 248-255
9. Shriver, Z., Hu, Y. N., Pojasek, K., and Sasiseskharan, R. (1998) Heparinase II from *Flavobacterium heparinum* - role of cysteine in enzymatic activity as probed by chemical, modification and site-directed mutagenesis. *J. Biol. Chem.* **273**, 22904-22912
10. Shriver, Z., Hu, Y., and Sasisekharan, R. (1998) Heparinase II from *Flavobacterium heparinum* - role of histidine residues in enzymatic activity as probed by chemical modification and site-directed mutagenesis. *J. Biol. Chem.* **273**, 10160-10167
11. Su, H. S., Blain, F., Musil, R. A., Zimmermann, J. J. F., Gu, K. F., and Bennett, D. C. (1996) Isolation and expression in *Escherichia coli* of *hepB* and *hepC*, genes coding for the glycosaminoglycan-degrading enzymes heparinase II and heparinase III, respectively, from

Flavobacterium heparinum. *Appl. Environ. Microbiol.* **62**, 2723-2734
12. Godavarti, R., Davis, M., Venkataraman, G., Cooney, C., Langer, R., and Sasisekharan, R. (1996) Heparinase III from *Flavobacterium heparinum*: cloning and recombinant expression in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **225**, 751-758
13. Pojasek, K., Shriver, Z., Hu, Y. N., and Sasisekharan, R. (2000) Histidine 295 and histidine 510 are crucial for the enzymatic degradation of heparan sulfate by heparinase III. *Biochemistry* **39**, 4012-4019
14. Han, Y. H., Garron, M. L., Kim, H. Y., Kim, W. S., Zhang, Z. Q., Ryu, K. S., Shaya, D., Xiao, Z. P., Cheong, C., Kim, Y. S., Linhardt, R. J., Jeon, Y. H., and Cygler, M. (2009) Structural snapshots of heparin depolymerization by heparin lyase I. *J. Biol. Chem.* **284**, 34019-34027
15. Shaya, D., Tocilj, A., Li, Y. G., Myette, J., Venkataraman, G., Sasisekharan, R., and Cygler, M. (2006) Crystal structure of heparinase II from *Pedobacter heparinus* and its complex with a disaccharide product. *J. Biol. Chem.* **281**, 15525-15535
16. Dong, W., Lu, W. Q., McKeehan, W. L., Luo, Y. D., and Ye, S. (2012) Structural basis of heparan sulfate-specific degradation by heparinase III. *Protein Cell* **3**, 950-961
17. Hashimoto, W., Kobayashi, E., Nankai, H., Sato, N., Miya, T., Kawai, S., and Murata, K. (1999) Unsaturated glucuronyl hydrolase of *Bacillus* sp. GL1: novel enzyme prerequisite for metabolism of unsaturated oligosaccharides produced by polysaccharide lyases. *Arch. Biochem. Biophys.* **368**, 367-374
18. Glaser, P., Rusniok, C., Buchrieser, C., Chevalier, F., Frangeul, L., Msadek, T., Zouine, M., Couvé, E., Lalioui, L., Poyart, C., Trieu-Cuot, P., and Kunst, F. (2002) Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Mol. Microbiol.* **45**, 1499-1513
19. Brinkkotter, A., Kloss, H., Alpert, C. A., and Lengeler, J. W. (2000) Pathways for the utilization of *N*-acetyl-galactosamine and galactosamine in *Escherichia coli*. *Mol. Microbiol.* **37**, 125-135
20. Marion, C., Stewart, J. M., Tazi, M. F., Burnaugh, A. M., Linke, C. M., Woodiga, S. A., and King, S. J. (2012) *Streptococcus pneumoniae* can utilize multiple sources of hyaluronic acid for growth. *Infect. Immun.* **80**, 1390-1398
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
22. Doublié, S., and Carter, C. W. J. (1992) Crystallization of Nucleic Acids and Proteins: A Practical Approach 2nd Ed., Oxford University Press, Oxford, UK
23. Laemmli, U. K. (1970) Cleavage of structural proteins during assembly of head of bacteriophage T4. *Nature* **227**, 680-685
24. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326
25. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213-221
26. Emsley, P., and Cowtan, K. (2004) *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 2126-2132
27. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* **53**, 240-255

28. Lovell, S. C., Davis, I. W., Adrendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by C^{α} geometry: Φ, Ψ and C^{β} deviation. *Proteins: Struct. Funct. Bioinform.* **50**, 437-450
29. DeLano, W. L. (2004) *The PyMOL Molecular Graphics System*. DeLano Scientific LLC, San Carlos, CA
30. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The protein data bank. *Nucleic Acids Res.* **28**, 235-242
31. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
32. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212-2221
33. Hashimoto, W., Maesaka, K., Sato, N., Kimura, S., Yamamoto, K., Kumagai, H., and Murata, K. (1997) Microbial system for polysaccharide depolymerization: enzymatic route for gellan depolymerization by *Bacillus* sp. GL1. *Arch. Biochem. Biophys.* **339**, 17-23
34. Ramachandran, G. N., and Sasisekharan, V. (1968) Conformation of polypeptides and proteins. *Adv. Protein Chem.* **23**, 283-437
35. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138
36. Ochiai, A., Yamasaki, M., Mikami, B., Hashimoto, W., and Murata, K. (2010) Crystal structure of exotype alginate lyase Atu3025 from *Agrobacterium tumefaciens*. *J. Biol. Chem.* **285**, 24519-24528
37. Garron, M. L., and Cygler, M. (2010) Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. *Glycobiology* **20**, 1547-1573
38. Shaya, D., Zhao, W. J., Garron, M. L., Xiao, Z. P., Cui, Q. Z., Zhang, Z. Q., Sulea, T., Linhardt, R. J., and Cygler, M. (2010) Catalytic mechanism of heparinase II Investigated by site-directed mutagenesis and the crystal structure with its substrate. *J. Biol. Chem.* **285**, 20051-20061
39. Hyun, Y. J., Lee, J. H., and Kim, D. H. (2010) Cloning, overexpression, and characterization of recombinant heparinase III from *Bacteroides stercoris* HJ-15. *Appl. Microbiol. Biotechnol.* **86**, 879-890
40. Ernst, S., Langer, R., Cooney, C. L., and Sasisekharan, R. (1995) Enzymatic degradation of glycosaminoglycans. *Crit. Rev. Biochem. Mol. Biol.* **30**, 387-444
41. Han, C., Spring, S., Lapidus, A., Del Rio, T. G., Tice, H., Copeland, A., Cheng, J. F., Lucas, S., Chen, F., Nolan, M., Bruce, D., Goodwin, L., Pitluck, S., Ivanova, N., Mavromatis, K., Mikhailova, N., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y. J., Jeffries, C. C., Saunders, E., Chertkov, O., Brettin, T., Goker, M., Rohde, M., Bristow, J., Eisen, J. A., Markowitz, V., Hugenholtz, P., Kyrpides, N. C., Klenk, H. P., and Detter, J. C. (2009) Complete genome sequence of *Pedobacter heparinus* type strain (HIM 762-3(T)). *Stand. Genomic Sci.* **1**, 54-62
42. Kankainen, M., Paulin, L., Tynkkynen, S., von Ossowski, I., Reunanen, J., Partanen, P., Satokari, R., Vesterlund, S., Hendrickx, A. P. A., Lebeer, S., De Keersmaecker, S. C. J., Vanderleyden, J., Hamalainen, T., Laukkanen, S., Salovuori, N., Ritari, J., Alatalo, E., Korpela, R., Mattila-Sandholm, T., Lassig, A., Hatakka, K., Kinnunen, K. T., Karjalainen, H., Saxelin, M., Laakso, K., Surakka, A., Palva, A., Salusjarvi, T., Auvinen, P., and de Vos, W. M. (2009) Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human-mucus binding protein. *Proc. Natl. Acad. Sci. USA* **106**, 17193-17198

SECTION 2

Recognition mechanism of bacterial unsaturated glucuronyl hydrolase for heparin disaccharides

UGL recognizes a double bond in ΔGlcUA and triggers the hydration of C-5 atom (1,2). Because ΔGlcUA is a component of all unsaturated GAG oligosaccharides produced by polysaccharide lyases, UGLs are essential for the complete degradation of GAGs. To date, all bacterial UGLs are classified as members of the GH-88 family. However, different substrate specificities accommodate structural diversities of unsaturated GAG oligosaccharides with different sugar residues, glycosidic bonds, and degrees of sulfation. For example, streptococcal UGLs and BacillusUGL prefer sulfated and unsulfated unsaturated disaccharides with 1,3-glycosidic bonds, respectively (CHAPTER I) (3). In contrast, Phep_2830, the UGL of *P. heparinus* degrades only unsaturated heparin disaccharides with 1,4-glycosidic bonds (4).

Recent studies have focused on the physiological function and structure of UGLs. Peculiar mechanisms of UGL catalysis are demonstrated by using artificial substrates (2). Disruption of UGL gene leads to reduced upper respiratory tract colonization by *S. pneumoniae* (5). Specific inhibitors of UGL are therefore expected to provide anti-bacterial drugs with no side effects. In addition, inducible mRNA expression of the streptococcal enzyme in the presence of GAG was demonstrated in CHAPTER I and structural determinants of the preference of streptococcal UGL for sulfated substrates with 1,3-glycosidic bonds was identified in CHAPTER II. Moreover, HepC depolymerizing heparan sulfate containing 1,4-glycosidic bonds was structurally analyzed in SECTION 1 of CHAPTER III. However, the UGL recognition mechanism for 1,4-glycosidic bond-type substrates from heparin and heparan sulfate remains unknown. Degradation of heparin and heparan sulfate with 1,4-glycosidic bonds is also considered important for bacterial adherence and invasion to host cells because heparin, heparan sulfate, hyaluronan, and chondroitin sulfate are constituents of mammalian extracellular matrices (6). HepC homologous gene is also present as a component of UGL gene cluster (SECTION 1 of CHAPTER III)

This SECTION deals with structure determination of *P. heparinus* UGL specific for unsaturated disaccharides with 1,4-glycosidic bonds and its binding modes to these substrates by X-ray crystallography, docking simulation, site-directed mutagenesis, and enzyme kinetics.

Materials and Methods

Materials. Unsaturated chondroitin disaccharides were purchased from Seikagaku Biobusiness. Unsaturated heparin disaccharides were purchased from Iduron and Sigma-Aldrich. Restriction endonucleases and DNA-modifying enzymes were purchased from Toyobo. All other analytical grade chemicals were obtained from commercial sources.

Overexpression. Overexpression systems for *P. heparinus* UGLs (Phep_2238, Phep_2649, and Phep_2830) were constructed in *E. coli* cells as follows. Three *P. heparinus* UGL genes were amplified from the genome of *P. heparinus* by PCR. PCR was performed using *P. heparinus* genomic DNA as a template, synthetic oligonucleotide primers, and KOD-FX-neo polymerase (Toyobo). The sequences of oligonucleotides with *Nde*I and *Xho*I sites added to their 5' regions for Phep_2238, Phep_2649, and Phep_2830 are shown in Table 3-2-1. The PCR conditions were as follows: 94°C for 2 min followed by 30 cycles of 98°C for 10 sec, 55°C for 30 sec, and 68°C for 30 sec. PCR products were ligated with *Hin*cII-digested pUC119 using Ligation High Ver. 2 (Toyobo) in accordance with the manufacturer's protocol, and the resulting plasmids were digested with *Nde*I and *Xho*I to isolate UGL gene fragments. Phep_2238 and Phep_2649 DNA fragments were ligated with *Nde*I- and *Xho*I-digested pET21b vectors, and the Phep_2830 DNA fragment was ligated into *Nde*I- and *Xho*I-digested pCold IV vector (Takara Bio) using Ligation High Ver. 2. The resulting plasmids were designated pET21b-Phep_2238, pET21b-Phep_2649, and pCold IV-Phep_2830, respectively. To overexpress Phep_2238, Phep_2649, and Phep_2830, *E. coli* strain BL21(DE3) cells were transformed with pET21b-Phep_2238 and pET21b-Phep_2649, and *E. coli* strain Rosetta-gami B cells (Novagen) were transformed with pCold IV-Phep_2830.

Microorganisms and Culture Conditions. To express Phep_2238 or Phep_2649, *E. coli* strain BL21(DE3) cells harboring pET21b-Phep_2238 or pET21b-Phep_2649 plasmids were cultured at 37°C in Luria broth (Sigma-Aldrich) supplemented with sodium ampicillin (100 μg ml^{-1}). When turbidity at 600 nm reached 0.3–0.7, IPTG was added to the culture to a final concentration of 0.1 mM, and the cells were further cultured at 16°C for 44 h. To overexpress Phep_2830, *E. coli* strain Rosetta-gami B cells harboring pCold IV-Phep_2830 were cultured at 37°C in Luria broth supplemented with sodium ampicillin (100 μg ml^{-1}) and chloramphenicol (33 μg ml^{-1}). When turbidity at 600 nm reached 1–1.2, the cells were cooled to 15°C with ice water, IPTG was added to the culture to a final concentration of 0.4 mM, and the cells were further cultured at 15°C for 44 h.

Purification. SagUGL, SpnUGL, SpyUGL, SagUGL mutant S368G (CHAPTER II), and BacillusUGL were expressed and purified as described in CHAPTERs I and II. *E. coli* cells

harboring pET21b-Phep_2238, pET21b-Phep_2649, or pCold IV-Phep_2830 plasmids were collected by centrifugation at 6,700 × *g* and 4°C for 10 min. Cells harboring pET21b-Phep_2238 or pET21b-Phep_2649 were resuspended in 20 mM KPB (pH 6.0), and the cells harboring pCold IV-Phep_2830 were resuspended in a solution containing 20 mM Tris–HCl (pH 7.5), 1 mM dithiothreitol, and 1 mM EDTA. Cells were ultrasonically disrupted (Insonator model 201M) at 9 kHz and 0°C for 10 min, and the supernatant obtained by centrifugation at 28,000 × *g* and 4°C for 20 min was used as a cell extract. Phep_2238 and Phep_2649 were purified by cation exchange chromatography (TOYOPEARL CM-650) followed by gel filtration chromatography (HiLoad 16/60 Superdex 75 pg). Phep_2830 was purified by anion exchange chromatography (HiLoad 16/10 Q Sepharose, GE Healthcare) followed by gel filtration chromatography (HiLoad 16/60 Superdex 75 pg). Purity was assessed using SDS-PAGE (7).

Table 3-2-1. Primers

Primers for molecular cloning		
Phep_2238	forward	5′-GGCATATGAAAAGAAACCTGATTTTAAAAA-3′
	reverse	5′-CCCTCGAGTTATTTCTGTAGTTTTTTATAACGC-3′
Phep_2649	forward	5′-GGCATATGAACATTATCAAGTCAAGTCC-3′
	reverse	5′-CCCTCGAGTCATAAACCTCTTTCTTTTA-3′
Phep_2830	forward	5′-GGCATATGAAATCACTACTCAGTGCGTTTG-3′
	reverse	5′-CCCTCGAGTTAAGACTGATTAATTGTTTTC-3′
Primers for site-directed mutagenesis		
S365A	sense	5′-TATTTATCTACGCATTATGCCGCACAGGGAAACCACCGTTTATTT-3′
	antisense	5′-AAATAAACGGTGGTTTCCCTGTGCGGCATAATGCGTAGATAAATA-3′
S365G	sense	5′-TTATCTACGCATTATGCCGAAGCGGGAAACCACCGTTTATTTGAA-3′
	antisense	5′-TTCAAATAAACGGTGGTTTCCCGCTTCGGCATAATGCGTAGATAA-3′
S368G	sense	5′-ACGCATTATGCCGAACAGGGAGCCCACCGTTTATTTGAAGCCCAA-3′
	antisense	5′-TTGGGCTTCAAATAAACGGTGGGCTCCCTGTTCGGCATAATGCGT-3′
K370A	sense	5′-CATTATGCCGAACAGGGAAACGCCCGTTTATTTGAAGCCCAACGC-3′
	antisense	5′-GCGTTGGGCTTCAAATAAACGGGCGTTTCCCTGTTCGGCATAATG-3′
D182N	sense	5′-CAGTTTGAACTTTCACCAATTTTCCATGTAGCTGCCATCGATATC-3′
	antisense	5′-GATATCGATGGCAGCTACATGGAAAATTGGTGAAAGTTCAAACTG-3′
R57A	sense	5′-ACCCCTATGTTTGGAGATTCAGCGATTACAGATAAAAATTTCAGG-3′
	antisense	5′-CCTGAAATTTTTATCTGTAATCGCTGAATCTCCAAACATAGGGGT-3′
F164A	sense	5′-AAAGCCAGTCCTCCCGGAGAAGCTCATGCCCAGCCGGATAACGGG-3′
	antisense	5′-CCCGTTATCCGGCTGGGCATGAGCTTCTCCGGGAGGACTGGCTTT-3′
D174N	sense	5′-GCCAGTCCTCCCGGAGAATTTGCTGCCCAGCCGGATAACGGGACT-3′
	antisense	5′-AGTCCCGTTATCCGGCTGGGCAGCAAATTCTCCGGGAGGACTGGC-3′
G362Y	sense	5′-CCAGACGCCGGGGTATTTGTGTTTAGCGGCGACGAAGCCATCATG-3′
	antisense	5′-CATGATGGCTTCGTCGCCGCTAAACACAAATACCCCGGCGTCTGG-3′

Underlines in primers for molecular cloning and site-directed mutagenesis indicate the restriction and mutation sites, respectively.

Enzyme Assay. Reactions of SagUGL, SpnUGL, SpyUGL, and BacillusUGL were conducted at 30°C in 500 µl solutions of 20 mM Tris–HCl (pH 7.5), 0.2 mM substrate, and enzyme. Reactions of Phep_2238, Phep_2649, and Phep_2830 were conducted at 30°C in 500 µl solutions containing 100 mM KPB (pH 6.0), 0.2 mM substrate, and enzyme. Concentrations and pH of each buffer were adopted in accordance with CHAPTER II and previous reports (4). Enzyme activity was measured by monitoring decreases in absorbance at 235 nm, which corresponded to the loss of substrate C–C double bonds.

Unsaturated chondroitin disaccharides were used as substrates, and their molar absorption coefficients at 235 nm [ε_{235} (M^{-1} cm^{-1})] were as follows: CΔ0S, ε_{235} = 4,800; CΔ4S, ε_{235} = 4,800; and CΔ6S, ε_{235} = 4,800. Unsaturated heparin disaccharides were also used as substrates, and their ε_{235} (M^{-1} cm^{-1}) were as follows: HΔNAc0S, ε_{235} = 4,524; HΔNS, ε_{235} = 6,600; HΔ6S, ε_{235} = 4,826; HΔNAc6S, ε_{235} = 4,300; HΔNS6S, ε_{235} = 6,075; HΔ2'SNS, ε_{235} = 4,433 (4).

Kinetic parameters of SagUGLs (wild-type, S365A, S365G, S368A, S368G, and K370A), BacillusUGL, Phep_2238, Phep_2649, and Phep_2830s (wild-type and F164A) with various unsaturated GAG disaccharides were determined using assays of UGL activity at 30°C, in which decreases in substrate absorbance at 235 nm were monitored. Reaction mixtures of SagUGLs and BacillusUGL contained substrate, 20 mM Tris–HCl (pH 7.5), and enzyme. Reaction mixtures for Phep_2238, Phep_2649, and Phep_2830s contained substrate, 100 mM KPB (pH 6.0), and enzyme. Substrate concentrations ranges were fixed at 0.05–1.0 mM because the absorbance at 235 nm of the substrate at over 1.0 mM exceeds measurement limitations on the spectrometer. Michaelis–Menten constants (K_m) and turnover numbers (k_{cat}) were calculated by fitting the data using the KaleidaGraph software.

Docking Simulations. To perform docking analyses of UGLs and substrates, coordinates of SagUGL were used. The structure of ligand-free SagUGL (CHAPTER I, PDB ID: 2ZZR) was selected as a receptor model. Docking analyses were performed using the *AutoDock 4.2* program (8). The coordinates of CΔ6S were obtained from coordinates of SagUGL complexed with CΔ6S (CHAPTER II, PDB ID: 3ANK), and coordinates of unsaturated heparin disaccharides were obtained using *ACD/ChemSketch Freeware*, version 5.12 (Advanced Chemistry Development, Inc.) and the *OpenBabel* program (9). Asp-175 and Lys-370 of SagUGL were treated as flexible residues because Asp-175 of SagUGL was critical for catalysis and Lys-370 showed movement during interactions with CΔ6S (CHAPTERs I and II). The number of genetic algorithm runs was set to 20. Figures for protein structures and docking forms were prepared using the *PyMol* program (10).

Site-directed Mutagenesis. Three residues (Ser-365, Ser-368, and Lys-370) of SagUGL were substituted with Ala or Gly, Ala, and Ala, and the resulting mutants were designated S365A, S365G, S368A, and K370A, respectively. Asp-182 of Phep_2238, and Arg-57, Phe-164, Asp-174, and Gly-362 of Phep_2830 were also substituted with Asn, Ala, Ala, Asn, and Tyr, and the resulting mutants were designated D182N, R57A, F164A, D174N, and G362Y, respectively. With the exception of D174N and G362Y, UGL mutants were constructed using a QuikChange site-directed mutagenesis kit. D174N and G362Y were constructed using a KOD-Plus Mutagenesis Kit (Toyobo) in accordance with the manufacturer's protocol. The plasmids pET21b-SagUGL and pCold IV-Phep_2830 were used as PCR templates for SagUGL and Phep_2830, respectively, and synthetic oligonucleotides were used as sense and antisense primers. Sense and antisense primer sequences are shown in Table 3-2-1. PCR was performed using KOD-Plus-Neo polymerase (Toyobo). Mutations were confirmed using the dideoxy-chain termination method with the automated DNA sequencer model 3730xl (11). *E. coli* strain HMS174(DE3) cells were transformed with SagUGL mutant plasmids. Expression and purification of SagUGL mutants were conducted using the procedure for wild-type SagUGL as described in CHAPTER II. Phep_2830 mutant plasmids were digested with *Nde*I and *Xho*I, and the resulting DNA fragments were ligated with *Nde*I and *Xho*I-digested pCold II vector (Takara Bio) to express Phep_2830 mutant proteins with N-terminal histidine tags for purification. N-terminal histidine-tagged Phep_2830 reportedly exhibits comparable enzyme activity to that of the histidine tag-free enzyme (4). The Phep_2830 mutants R57A, F164A, and G362Y were purified to almost homogeneity by affinity chromatography (TALON).

X-ray Crystallography. To determine the three-dimensional structure of Phep_2830, the purified enzyme was crystallized using sitting drop vapor diffusion. Solutions containing 1 μl proteins (10 mg ml^{-1}) in 20 mM Tris–HCl (pH 7.5), 1 mM dithiothreitol, and 1 mM EDTA were mixed with equal volumes of reservoir solution containing 25% polyethylene glycol 6,000, 0.1 M HEPES (pH 7.5), and 0.1 M LiCl. Protein solutions were then incubated at 20°C, and single crystals were grown for about 2 months. The crystal was flash-cooled under a cold nitrogen stream. Diffraction data were collected at λ = 1.0000 Å using an ADSC Q315 detector at the BL38B1 station of SPring-8. The data were processed using the *HKL2000* program (12). Molecular replacements for structure determinations were conducted using the *Molrep* program supplied in the *CCP4* program package with coordinates of SagUGL (PDB ID: 2ZZR) as an initial model (13,14). Structure refinement was conducted with the *Phenix.refine* program supplied in the *PHENIX*

program package (15,16). Randomly selected 5% reflections were excluded from refinement and were used to calculate R_{free}. After each cycle of refinement, the model was manually adjusted using the *Coot* program (17). The final model quality was checked using the *PROCHECK* program (18). Superpositioning of protein models and calculation of their r.m.s.d. were conducted using the *LSQKAB* program supplied with the *CCP4* program package (19). The atomic coordinates and structure factors of Phep_2830 (ID: 3WIW) have been deposited in the PDB (http://wwpdb.org/).

Results and Discussion

Enzyme Activity of Streptococcal UGLs toward Unsaturated Heparin Disaccharides

Three streptococcal UGLs (SagUGL, SpnUGL, and SpyUGL) were purified to homogeneity, and their enzyme activities (U mg^{-1}) toward a variety of unsaturated heparin disaccharides were measured (Table 3-2-2). All three streptococcal UGLs degraded HΔNS with greater efficiency than the other unsaturated heparin disaccharides. Kinetic parameters of SagUGL toward HΔNS were determined as follows: K_m, 1.9 mM; k_{cat}, 3.9 s^{-1}. SagUGL is known to exhibit the highest enzymatic activity toward CΔ6S with 1,3-glycosidic bond (CHAPTER II) as follows: K_m, 0.10 mM; k_{cat}, 10 s^{-1}, indicating that the enzyme activity of SagUGL toward unsaturated heparin disaccharides was lower than that toward unsaturated chondroitin disaccharides. Indeed, k_{cat} K_m^{-1} with HΔNS was 50-fold lower than with CΔ6S (CHAPTER II). The enzyme activities of BacillusUGL with each unsaturated heparin disaccharide were also measured. BacillusUGL exhibited the highest enzyme activity toward HΔNAc0S distinct from streptococcal UGLs (Table 3-2-2). BacillusUGL also demonstrated specificity for CΔ0S (3). These results and previous data (CHAPTER I) (3) demonstrate that streptococcal UGLs exhibit high enzyme activity toward disaccharides with specific sulfate groups, whereas BacillusUGL preferentially degrades unsulfated substrates.

Recognition of Unsaturated Heparin Disaccharides by Streptococcal UGL

Structure determinations of SagUGL in complex with HΔNS were difficult, possibly due to its lower affinity for this disaccharide. Thus, to elucidate the binding modes of unsaturated heparin disaccharide with SagUGL, the structure of SagUGL in complex with HΔNS was predicted using the docking simulation program *AutoDock*. To evaluate the suitability of the *AutoDock* program, docking simulations were performed with coordinates of ligand-free SagUGL determined in CHAPTER I (PDB ID: 2ZZR) and CΔ6S. The simulated complex structure with the lowest binding energy was almost identical to the crystal structure of SagUGL/CΔ6S except that Lys-370 forms hydrogen bond with a sulfate group determined in CHAPTER II (PDB ID: 3ANK; Fig. 3-2-1a, b),

indicating that the *AutoDock* docking simulation accurately determined the substrate-bound UGL structure. Thus, the binding mode of HΔNS to SagUGL was calculated using *AutoDock* program, indicating the accommodation of HΔNS in the active site of SagUGL (Fig. 3-2-1c). The direction of the pyranose ring of GlcN toward ΔGlcUA of HΔNS opposed that of GalNAc in CΔ6S (Fig. 3-2-1f). Due to these differences in direction, the amino group of GlcN in HΔNS is located at a similar position to the C-6 of GalNAc in CΔ6S, and the position of a sulfate group in HΔNS corresponds to that of a sulfate group in CΔ6S. The directions of HΔNS pyranose rings indicated by docking simulations were similar to those of heparin in solution, which were determined by X-ray scattering (20). This docking simulation suggests that Ser-365, Ser-368, and Lys-370 residues of SagUGL form hydrogen bonds with the sulfate group of HΔNS. To confirm the functions of Ser-365, Ser-368, and Lys-370 of SagUGL, the SagUGL mutants S365A, S365G, S368A, S368G, and K370A were purified and assayed. Kinetic parameters of each mutant were determined and compared with those of wild-type SagUGL (Table 3-2-3). K_m values of S365A, S365G, S368A, and S368G for HΔNS were higher than that of the wild-type enzyme, whereas that of K370A was lower. The k_{cat} values of S365G, S368G, and in particular K370A for HΔNS were also decreased (k_{cat} of K370A was ~50-fold lower than that of wild-type). This mutant analysis shows that Ser-365 and Ser-368 of SagUGL are involved in binding to HΔNS, whereas Lys-370 enhances the reaction efficiency of SagUGL rather than the binding affinity for HΔNS. In contrast, although the binding modes of the other unsaturated heparin disaccharides to SagUGL were calculated, no structures of substrate-bound SagUGL were obtained. These docking simulations and enzyme assays of SagUGL (Table 3-2-2) suggest that, with the exception of HΔNS, its affinity toward unsaturated heparin disaccharides is also remarkably low.

Table 3-2-2. Substrate specificity of bacterial UGLs

	Specific activity (U mg^{-1})						
	SagUGL	SpnUGL	SpyUGL	Bacillus UGL	Phep_2238	Phep_2649	Phep_2830
HΔNAc0S	0.016	0.0055	0.0033	5.3	12	0.30	15
HΔNS	0.77	0.14	0.10	1.8	11	1.1	7.1
HΔ6S	0.0088	0.0018	0.00094	0.081	1.2	ND[a]	1.9
HΔNAc6S	0.035	0.0059	0.0066	0.016	7.9	ND	18
HΔNS6S	0.0080	0.0031	ND	0.0050	4.2	ND	4.5
HΔ2'SNS	0.0023	0.0024	ND	0.0095	ND	ND	ND
CΔ0S	0.72	0.99	0.63	10	2.6	0.70	ND
CΔ4S	0.0019	ND	0.0046	ND	0.31	ND	ND
CΔ6S	17	6.1	5.3	1.8	3.2	1.7	ND

[a] Not detected

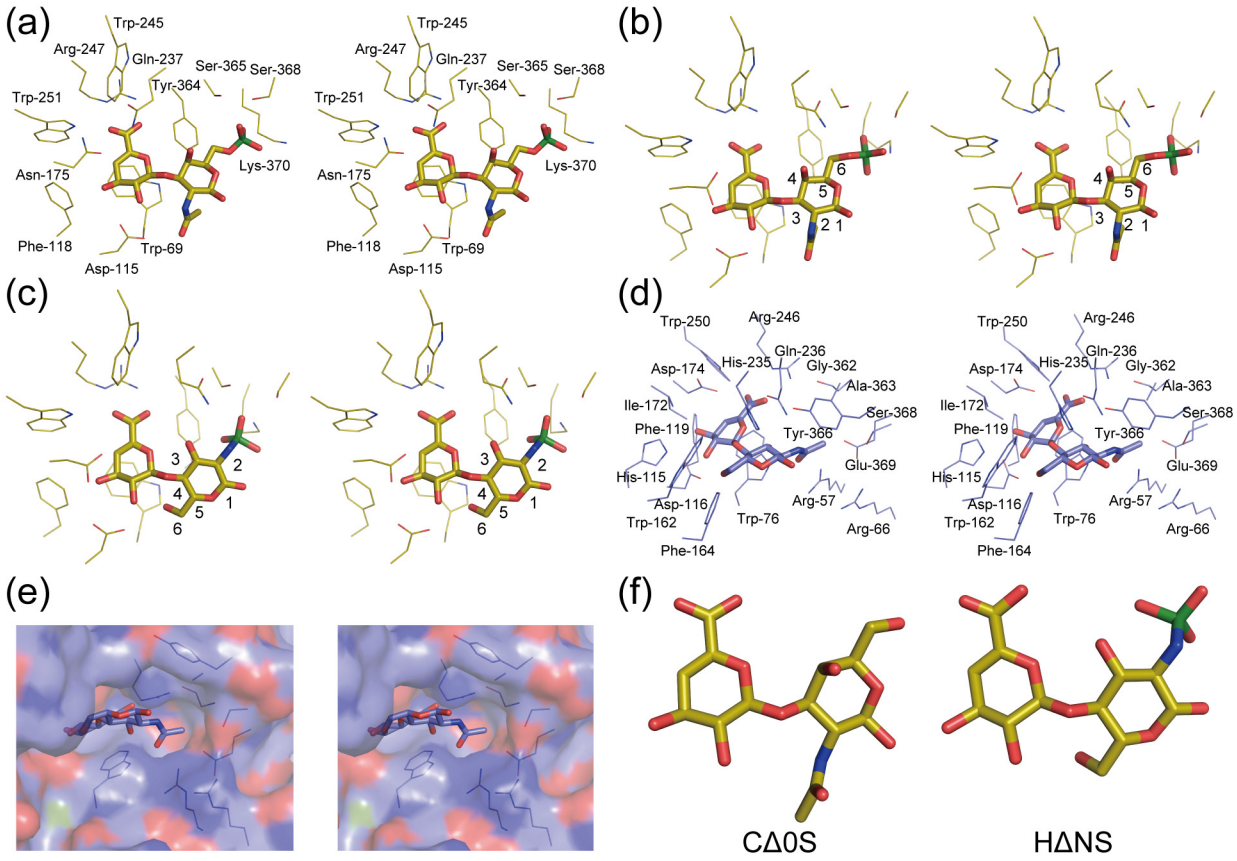

Figure 3-2-1. Binding mode of unsaturated disaccharides to bacterial UGLs. (a) Crystal structure of SagUGL in complex with CΔ6S determined in CHAPTER II (PDB ID: 3ANK). (b) An active site structure of SagUGL/CΔ6S. (c) An active site structure of SagUGL/HΔNS. (d) An active site structure of Phep_2830/HΔNAc0S. (e) Surface of the pocket recognizing acetyl group. (f) Directions of pyranose rings of an unsaturated chondroitin disaccharide and an unsaturated heparin disaccharide in complex with UGLs. (a)–(e) Stereo-diagram. (b)–(e) Structures estimated via docking simulations. Numbers of carbon atoms are shown (b) and (c). Atoms C of SagUGL, C of Phep_2830, O, N, and S are in *yellow*, *purple*, *red*, *blue*, and *green*, respectively.

Docking simulations of HΔNS with SagUGL, site-directed mutagenesis, and kinetic analyses demonstrate that Ser-365 and Ser-368 residues of SagUGL contribute to recognition of the sulfate group of HΔNS and that the Lys-370 residue is involved in degradation of the substrate. These amino acid residues constitute the motif S*XX*(S)*X*K, which plays an important role in the recognition of the sulfate group of CΔ6S (CHAPTER II). Streptococcal UGLs acted on unsaturated heparin disaccharides, although their enzyme activities toward these substrates were low. In addition, putative heparan sulfate lyase is encoded in the vicinity of the streptococcal UGL gene, suggesting that streptococcal UGLs contribute to the degradation of heparan sulfate.

Table 3-2-3. Kinetic parameters of SagUGL mutants for HΔNS

	K_m (mM)	k_{cat} (s^{-1})	$k_{cat}\,K_m^{-1}$ ($mM^{-1}\,s^{-1}$)
wild-type	1.9 ± 0.2	3.8 ± 0.4	2.0 ± 0.3
S365A	3.4 ± 1.3	1.9 ± 0.6	0.56 ± 0.14
S365G	5.2 ± 2.7	0.76 ± 0.35	0.15 ± 0.07
S368A	3.0 ± 1.0	6.0 ± 1.7	2.0 ± 0.4
S368G	3.4 ± 0.7	2.8 ± 0.5	0.83 ± 0.05
K370A	0.86 ± 0.29	0.075 ± 0.014	0.086 ± 0.013

Substrate Specificity of *P. heparinus* UGLs

The structural determinants of SagUGL recognition of HΔNS sulfate groups were calculated using docking simulations. However, the intrinsic UGL-binding mechanism of substrates containing 1,4-glycosidic bonds remains to be clarified. Therefore, the mechanisms behind UGL recognition of unsaturated heparin disaccharides were analyzed using high affinity UGL interactions with unsaturated heparin disaccharides. One (Phep_2830) of the UGLs of *P. heparinus* is known to specifically degrade substrates containing 1,4-glycosidic bonds. Thirteen UGL genes of *P. heparinus* were assigned by complete genome sequences (21) and the three UGLs Phep_2238, Phep_2649, and Phep_2830 were overexpressed in *E. coli* cells and purified to homogeneity and assayed for activity (Table 3-2-2).

Phep_2238 degraded HΔNAc0S and HΔNS more efficiently than other unsaturated heparin disaccharides. Moreover, Phep_2238 exhibited comparably lower enzyme activity toward substrates with 1,3-glycosidic bonds, such as CΔ0S, CΔ4S, and CΔ6S, than those with 1,4-glycosidic bonds. The kinetic parameters K_m and k_{cat} of Phep_2238 toward HΔNS were 0.073 mM and 2.3 s^{-1}, respectively. Phep_2649 preferred CΔ6S but also degraded CΔ0S, HΔNAc0S, and HΔNS, although its specific activity was lower than that of the streptococcal UGLs, BacillusUGL, Phep_2238, and Phep_2830. K_m and k_{cat} values of Phep_2649 toward HΔNS were 0.052 mM and 0.17 s^{-1}, respectively. Although a UGL that degrades unsaturated chondroitin disaccharide (CΔ6S) was previously isolated as a 1,3-glycuronidase from *P. heparinus* (22), this enzyme differs from Phep_2238 and Phep_2649 in amino acid composition, isoelectric point, and molecular weight. In contrast, Phep_2830 exhibited enzyme activity toward only unsaturated heparin disaccharides, as described previously (4). K_m and k_{cat} values of Phep_2830 toward HΔNAc6S were 0.029 mM and 16 s^{-1}, respectively.

Crystal Structure of Phep_2830.

Because Phep_2238 and Phep_2830 showed high enzyme activity toward various unsaturated

heparin disaccharides, these enzymes were crystallized to clarify the mechanisms by which UGL recognizes substrates with 1,4-glycosidic bonds based on tertiary structures. Phep_2830 was successfully crystallized, and X-ray diffraction data were collected. The crystal structure of Phep_2830 was determined at a resolution of 1.35 Å using molecular replacement with the SagUGL structure (CHAPTER I, PDB ID: 2ZZR) as an initial model. Data collection and model refinement statistics are summarized in Table 3-2-4. The final model contains one monomer enzyme from Gly-28 to Thr-398, a molecule of HEPES, and 422 water molecules. The N-terminal region from Met-1 to Asn-27 could not be assigned because the electron density map was too thin. Similar to SagUGL and BacillusUGL, the overall structure of Phep_2830 has a α_6/α_6-barrel architecture, and Phep_2830 contains twelve α helices and five β strands (Fig. 3-2-2a). The r.m.s.d. for all of the 336 C^{α} atoms between Phep_2830 and SagUGL was 1.7 Å, indicating that both have a common basic scaffold structure. In contrast, approximately ten C-terminal amino acid residues of Phep_2830 protrude forward from the outside of the protein. Previous study and CHAPTER I show that Asp-149 of BacillusUGL and Asp-175 of SagUGL act as critical catalysts (1). Thus, to investigate the catalytic mechanisms of *P. heparinus* UGLs, corresponding residues of Phep_2238 (Asp-182) and Phep_2830 (Asp-174) were substituted with Asn. Both mutant enzymes were inactive, indicating that the catalytic mechanisms of Phep_2238 and Phep_2830 are similar to those of SagUGL and BacillusUGL.

Active sites were structurally compared by superimposing coordinates of Phep_2830 on those of SagUGL (CHAPTER I, PDB ID: 2ZZR) and BacillusUGL/CΔ0S (PDB ID: 2AHG; Fig. 3-2-2b–d). Subsites were defined such that –n represents the nonreducing terminus, +n represents the reducing terminus, and cleavage occurs between these sites (23). The amino acid residues and their positions at subsite –1, which is the binding site of ΔGlcUA, were common to all three UGLs (Fig. 3-2-2b). However, the structures at subsite +1, which is the binding site of an amino sugar, differed significantly (Fig. 3-2-2c, d). In particular, Ser-365, Ser-368, and Lys-370 of SagUGL comprise the motif S*XX*(S)*X*K, which contributes to the recognition of a sulfate group (CHAPTER II). These three residues are not conserved in Phep_2830 and correspond to the Ala-363, Tyr-366, and Ser-368 residues of Phep_2830, respectively. These observations indicate that Phep_2830 recognition of substrate sulfate groups differs from that of SagUGL. Regarding hydrophobic amino acid residues forming a stacking interaction with an amino sugar at the subsite +1, the Tyr-338 residue of BacillusUGL (or the Tyr-364 residue of SagUGL), which is located around the C-6 position of GalNAc in BacillusUGL/CΔ0S is substituted with the Gly-362 residue in Phep_2830,

Table 3-2-4. Data collection and refinement statistics

Data collection	
Wavelength (Å)	1.0000
Space group	*C*2
Unit cell parameters (Å, deg)	a = 91.8, b = 51.0, c = 85.8 β = 115.7
Resolution limit (Å)	50.00–1.35 (1.40–1.35)
Measured reflections	368,700
Unique reflections	78,862
Redundancy	4.7 (3.9)
Completeness (%)	99.9 (100)
I/σ (I)	32.5 (3.72)
R_{merge}	0.091 (0.37)
Refinement	
R_{work}	0.132
R_{free}	0.160
No. of molecules/asymmetric unit	1
No. of non-hydrogen atoms	
Protein	3,244
HEPES	15
Water molecules	422
Average isotropic *B*-factor (Å^2)	
Protein	16.7
HEPES	30.1
Water molecules	27.5
r.m.s.d. from ideal	
Bond length (Å)	0.009
Bond angle (deg)	1.226
Ramachandran plot (%)	
Favored region	97.8
Allowed region	2.2
Outlier region	0

[a] Data on highest shells are in parenthesis.

whereas the Trp-134 residue of BacillusUGL (or the Trp-161 residue of SagUGL) is conserved in Phep_2830. Due to this difference, this large space in Phep_2830 occurs around Gly-362, and is distinct from that of SagUGL and BacillusUGL. As described above, the Phe-164 residue of Phep_2830 that corresponds to Glu-163 of SagUGL or Pro-136 of BacillusUGL also lies at the binding site of an amino sugar (Fig. 3-2-2c, d). The loop comprising amino acid residues 163–169 of Phep_2830, which is designated loop A and corresponds to residues 162–170 of SagUGL and 135–144 of BacillusUGL, covers the active site. However, these loops of both SagUGL and BacillusUGL are distal from those at the active site. Accordingly, the Phe-164 residue of Phep_2830 is proximal to the active site (Fig. 3-2-2d). These Phep_2830 specific amino acid residues may

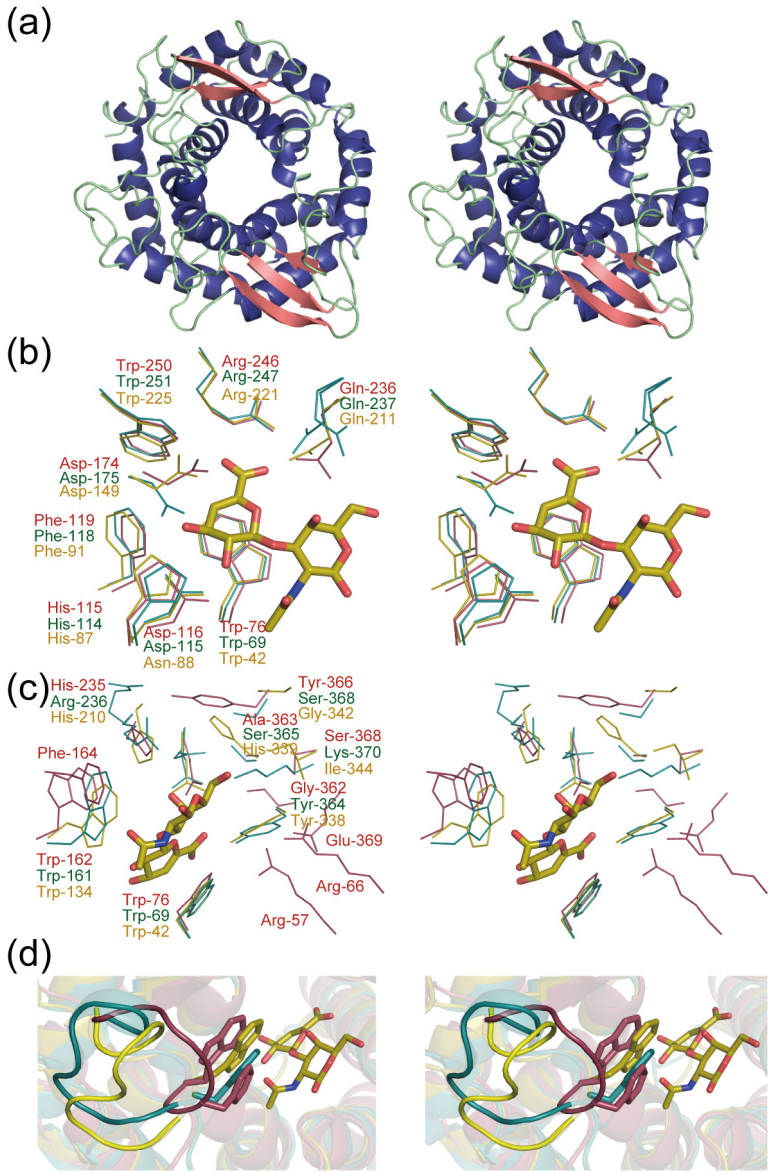

Figure 3-2-2. Crystal structure of Phep_2830 (stereo-diagram). (a) Overall structure of Phep_2830. *Blue*, *pink*, and *green* denote the secondary structure elements α-helices, β-sheets, and loops, respectively. (b) Amino acid residues of the three UGLs Phep_2830, SagUGL, and BacillusUGL/CΔ0S at subsite −1. (c) Amino acid residues of the three UGLs at subsite +1. (d) Each loop comprises amino acid residues 163–169 of Phep_2830, 162–170 of SagUGL, and 135–144 of BacillusUGL. Sticks indicate side chains of Trp-162 and Phe-164 of Phep_2830, Trp-161 of SagUGL, and Trp-134 of BacillusUGL. (b)–(d) Phep_2830, SagUGL (CHAPTER I, PDB ID: 2ZZR), and BacillusUGL in complex with CΔ0S (PDB ID: 2AHG) are indicated in *purple*, *cyan*, and *yellow*, respectively. Atoms O and N of CΔ0S are colored *red* and *blue*, respectively.

be involved in stacking interactions with substrates, especially with amino sugars. Indeed, whereas Arg-57, Arg-66, and Glu-369 residues of Phep_2830 are arranged at subsite +1, these are not conserved in the active sites of either SagUGL or BacillusUGL.

Binding Mode of Unsaturated Heparin Disaccharides to Phep_2830

To clarify the mechanism by which Phep_2830 recognizes unsaturated heparin disaccharides, preparation of complexes of Phep_2830 with these substrates was tried but failed. Thus, the binding modes of unsaturated heparin disaccharides (HΔNAc0S, HΔNS, HΔ6S, and HΔNAc6S) to Phep_2830 were estimated in a similar way to those of SagUGL using the *AutoDock* program. In these simulations, structures of HΔNAc0S, HΔNS, and HΔNAc6S-bound Phep_2830 were successfully calculated and that of the Phep_2830–HΔNAc0S complex is shown in Fig. 3-2-1d. Two common features were observed in the three docking structures that exhibited low binding energies. Similar to the docking simulation of SagUGL with HΔNS, the direction of the pyranose ring of GlcNAc (or GlcN sulfated at the *N* position) toward ΔGlcUA was opposed to that of GalNAc in complex with SagUGL/CΔ6S and BacillusUGL/CΔ0S (Fig. 3-2-1f). In addition, the acetyl groups of HΔNAc0S and HΔNAc6S (or the sulfate group of HΔNS) were predicted to be accommodated in the pocket-like structure comprising Arg-57, Arg-66, Trp-73, Gly-362, Ala-363, Tyr-366, Ser-368, and Glu-369 residues, which was designated the acetyl/sulfate group-binding pocket (Fig. 3-2-1d, e). No similar pocket-like structures were observed in SagUGL and BacillusUGL, which prefer substrates with 1,3-glycosidic bonds (Fig. 3-2-2c). The Trp-162 and Phe-164 residues in loop A of Phep_2830 (Fig. 3-2-2d) are located close to the C-6 position of an amino sugar of unsaturated heparin disaccharides. As described above, Trp-161, Trp-134, and Trp-162 residues of SagUGL, BacillusUGL, and Phep_2830, respectively, are situated at almost identical positions, whereas the residues of SagUGL and BacillusUGL that correspond to Phe-164 of Phep_2830 are distal from subsite +1.

To confirm that this pocket-like structure and loop A contribute to recognition of unsaturated heparin disaccharides, Arg-57, Phe-164, and Gly-362 residues of Phep_2830 were substituted with Ala, Ala, and Tyr, respectively. The resulting mutants R57A, F164A, and G362Y with N-terminal His-tags were expressed in *E. coli* and were purified using affinity columns. Although their expression was confirmed by SDS-PAGE (Fig. 3-2-3), two mutants of pocket-like structure (R57A and G362Y) exhibited no detectable enzyme activity with unsaturated heparin disaccharides as well as unsaturated chondroitin disaccharides. The K_m value (1.4 mM) of F164A toward HΔNAc6S was about 50 fold higher than that (0.029 mM) of the wild-type enzyme, whereas there was no significant difference in k_{cat} between wild-type enzyme (16 s^{-1}) and F164A (22 s^{-1}). These site-directed mutagenesis experiments indicated that the pocket-like structure and loop A play a significant role in recognizing substrates. In contrast, the absence of the acetyl group may have led

to low specific activity of Phep_2830 toward HΔ6S and hampered docking simulations of the complex of Phep_2830 with HΔ6S. Hence, substrates containing 1,4-glycosidic bonds are readily inserted into the acetyl/sulfate group-binding pocket of Phep_2830 through stacking interactions of loop A residues, Trp-162 and Phe-164, with an amino sugar. Simultaneously, Trp-162 and Phe-164 residues may inhibit binding of Phep_2830 with substrates containing 1,3-glycosidic bonds.

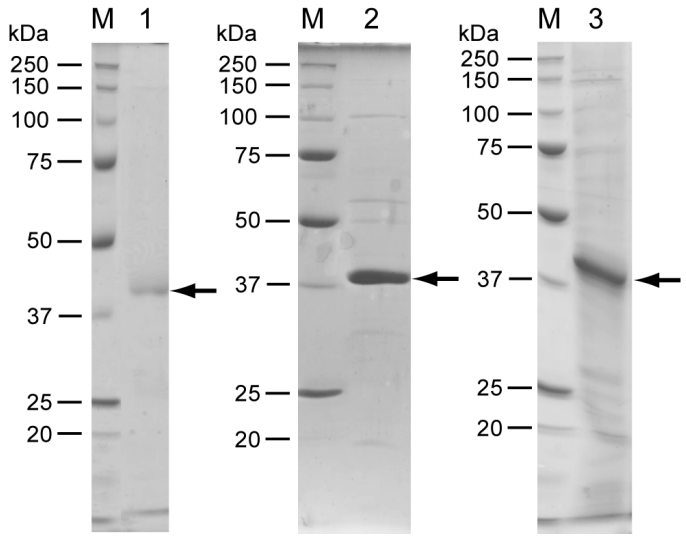

Figure 3-2-3. SDS-PAGE profile. *Lane M*, molecular weight standards (from top): synthetic polypeptides with molecular weights of 250,000, 150,000, 100,000, 75,000, 50,000, 37,000, 25,000, and 20,000; *lane 1*, R57A; *lane 2*, F164A; *lane 3*, G362Y.

Comparisons of the active site structures and docking simulations suggest that the acetyl/sulfate group-binding pocket comprising Arg-57, Arg-66, Trp-73, Gly-362, Ala-363, Tyr-366, Ser-368, and Glu-369, and the loop A including Trp-162 and Phe-164, are important for binding of Phep_2830 to unsaturated heparin disaccharides. Among these amino acid residues, Arg-57, Trp-73, Gly-362, Glu-369, and Trp-162 are conserved in Phep_2238 and Phep_2649 (Fig. 3-2-4). Moreover, Arg-66, Ala-363, Tyr-366, Ser-368, and Phe-164 of Phep_2830 correspond with Val-73, His-373, Gly-376, Ser-378, and Ser-171 of Phep_2238 and Arg-59, Ser-362, His-365, Asn-367, and Val-158 of Phep_2649, respectively. These differences in primary structure may lead to changes in pocket shapes but do not occupy the space. However, Ser-171 of Phep_2238 and Val-158 of Phep_2649, which correspond to Phe-164 of Phep_2830, do not form stacking interactions because they are smaller than Phe. Hence, these structural features may enable recognition of substrates with 1,3-glycosidic bonds by Phep_2238 and Phep_2649. Although the pocket-like structure of Phep_2649 was arranged with high probability, the activity of the enzyme for unsaturated heparin

disaccharides was lower than that of Phep_2238 and Phep_2830. In addition, the relative length of amino acid residues of Phep_2649 that correspond to the loop A may contribute to low enzyme activity.

Bacteroides species, human intestinal bacteria, are known to produce family PL-12, 13, and 21 heparinases depolymerizing heparin to unsaturated disaccharides through β-elimination reaction (24-28), whereas complete degradation of heparin to monosaccharides remains to be clarified. Four UGL-homologous proteins are encoded in the genomes of *B. stercoris* (BACSTE_02470, BACSTE_03202, BACSTE_03210, and BACSTE_03707) and *B. thetaiotaomicron* (BT_0146, BT_2913, BT_3348, and BT_4658). Arg-57 and Gly-362, key residues comprising the pocket-like structure of Phep_2830, are conserved in these bacteroides UGLs, suggesting that, similar to Phep_2238 and Phep_2830, bacteroides enzymes preferentially degrade unsaturated heparin disaccharides to constituent monosaccharides. Heparin is abundant in lung, ileum, and skin of mammals (29). The UGLs containing primary structure for recognition of unsaturated heparin disaccharides are expected to facilitate the indigenous bacteria, such as *Bacteroides*, to stay in the intestine. Consequently, structural determinants of Phep_2830 for 1,4-specificity found in this SECTION promotes a better understanding of the complete degradation of heparin by *Bacteroides* species and interaction between normal flora and host cells through GAG degradation.

In conclusion, tertiary and active site structures of UGLs that are specific for unsaturated heparin disaccharides containing 1,4-glycosidic bonds were determined for the first time. Comparisons of the active site structures, docking simulations, and site-directed mutagenesis experiments demonstrate the significance of the acetyl/sulfate group-binding pocket and the lid loop at subsite +1 in Phep_2830 for recognition of substrates with 1,4-glycosidic bonds.

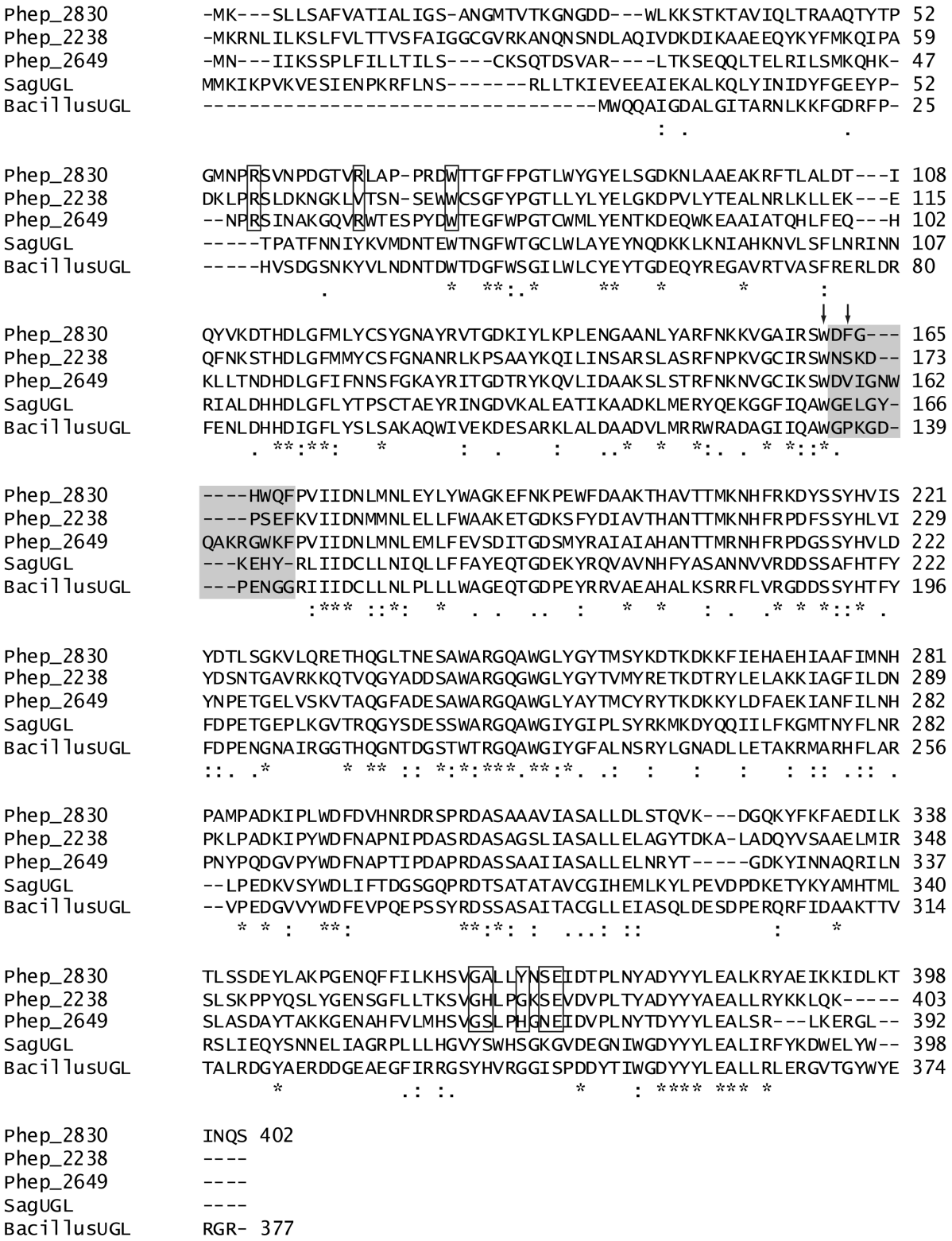

Figure 3-2-4. Multiple sequence alignments of UGLs. Primary structures of Phep_2238, Phep_2649, Phep_2830, SagUGL, and BacillusUGL were used for sequence alignments. Identical and similar amino acid residues among proteins are denoted by *asterisks* and *dots*, respectively. Amino acid residues of the acetyl/sulfate group-binding pocket are indicated in *boxes*. The residues comprising loop A are *shaded*. *Arrows* indicate positions of Trp-162 and Phe-164 of Phep_2830.

References

1. Itoh, T., Hashimoto, W., Mikami, B., and Murata, K. (2006) Crystal structure of unsaturated glucuronyl hydrolase complexed with substrate. *J. Biol. Chem.* **281**, 29807-29816
2. Jongkees, S. A. K., and Withers, S. G. (2011) Glycoside cleavage by a new mechanism in unsaturated glucuronyl hydrolases. *J. Am. Chem. Soc.* **133**, 19334-19337
3. Hashimoto, W., Kobayashi, E., Nankai, H., Sato, N., Miya, T., Kawai, S., and Murata, K. (1999) Unsaturated glucuronyl hydrolase of *Bacillus* sp. GL1: novel enzyme prerequisite for metabolism of unsaturated oligosaccharides produced by polysaccharide lyases. *Arch. Biochem. Biophys.* **368**, 367-374
4. Myette, J. R., Shriver, Z., Kiziltepe, T., McLean, M. W., Venkataraman, G., and Sasisekharan, R. (2002) Molecular cloning of the heparin/heparan sulfate Δ4,5 unsaturated glycuronidase from *Flavobacterium heparinum*, its recombinant expression in *Escherichia coli*, and biochemical determination of its unique substrate specificity. *Biochemistry* **41**, 7424-7434
5. Marion, C., Stewart, J. M., Tazi, M. F., Burnaugh, A. M., Linke, C. M., Woodiga, S. A., and King, S. J. (2012) *Streptococcus pneumoniae* can utilize multiple sources of hyaluronic acid for growth. *Infect. Immun.* **80**, 1390-1398
6. Sasisekharan, R., Raman, R., and Prabhakar, V. (2006) Glycomics approach to structure-function relationships of glycosaminoglycans. *Annu. Rev. Biomed. Eng.* **8**, 181-231
7. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
8. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785-2791
9. O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., and Hutchison, G. R. (2011) Open Babel: an open chemical toolbox. *J. Cheminform.* **3**, 33
10. DeLano, W. L. (2004) *The PyMOL Molecular Graphics System*. DeLano Scientific LLC, San Carlos, CA
11. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
12. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326
13. Vagin, A. A., and Isupov, M. N. (2001) Spherically averaged phased translation function and its application to the search for molecules and fragments in electron-density maps. *Acta Crystallogr. D. Biol. Crystallogr.* **57**, 1451-1456
14. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the *CCP4* suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235-242
15. Afonine, P. V., Grosse-Kunstleve, R. W., and Adams, P. D. (2005) The Phenix refinement framework. *CCP4 Newsletter* **42**, 8
16. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213-221
17. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of

Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486-501
18. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *PROCHECK* - a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291
19. Kabsch, W. (1976) Solution for best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. A* **32**, 922-923
20. Khan, S., Gor, J., Mulloy, B., and Perkins, S. J. (2010) Semi-rigid solution structures of heparin by constrained X-ray scattering modelling: new insight into heparin-protein complexes. *J. Mol. Biol.* **395**, 504-521
21. Han, C., Spring, S., Lapidus, A., Del Rio, T. G., Tice, H., Copeland, A., Cheng, J. F., Lucas, S., Chen, F., Nolan, M., Bruce, D., Goodwin, L., Pitluck, S., Ivanova, N., Mavromatis, K., Mikhailova, N., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y. J., Jeffries, C. C., Saunders, E., Chertkov, O., Brettin, T., Goker, M., Rohde, M., Bristow, J., Eisen, J. A., Markowitz, V., Hugenholtz, P., Kyrpides, N. C., Klenk, H. P., and Detter, J. C. (2009) Complete genome sequence of *Pedobacter heparinus* type strain (HIM 762-3(T)). *Stand. Genomic Sci.* **1**, 54-62
22. Gu, K. N., Linhardt, R. J., Laliberte, M., Gu, K. F., and Zimmermann, J. (1995) Purification, characterization and specificity of chondroitin lyases and glycuronidase from *Flavobacterium heparinum*. *Biochem. J.* **312**, 569-577
23. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557-559
24. Kim, W. S., Kim, B. T., Kim, D. H., and Kim, Y. S. (2004) Purification and characterization of heparin lyase I from *Bacteroides stercoris* HJ-15. *J. Biochem. Mol. Biol.* **37**, 684-690
25. Hyun, Y. J., Lee, J. H., and Kim, D. H. (2010) Cloning, overexpression, and characterization of recombinant heparinase III from *Bacteroides stercoris* HJ-15. *Appl. Microbiol. Biotechnol.* **86**, 879-890
26. Hyun, Y. J., Lee, K. S., and Kim, D. H. (2010) Cloning, expression and characterization of acharan sulfate-degrading heparin lyase II from *Bacteroides stercoris* HJ-15. *J. Appl. Microbiol.* **108**, 226-235
27. Luo, Y. D., Huang, X. Q., and McKeehan, W. L. (2007) High yield, purity and activity of soluble recombinant *Bacteroides thetaiotaomicron* GST-heparinase I from *Escherichia coli*. *Arch. Biochem. Biophys.* **460**, 17-24
28. Dong, W., Lu, W. Q., McKeehan, W. L., Luo, Y. D., and Ye, S. (2012) Structural basis of heparan sulfate-specific degradation by heparinase III. *Protein Cell* **3**, 950-961
29. Nader, H. B., Takahashi, H. K., Straus, A. H., and Dietrich, C. P. (1980) Selective distribution of the heparin in mammals - conspicuous presence of heparin in lymphoid tissues. *Biochim. Biophys. Acta* **627**, 40-48

CONCLUSION

1. The GAG-related genes for hyaluronate lyase, UGL, and phosphotransferase system for amino sugar, a component of GAGs, are assembled into a cluster (UGL genetic cluster) in the genome of pyogenic and hemolytic streptococci such as *S. agalactiae*, *S. pneumoniae*, and *S. pyogenes*. The UGL genetic cluster of *S agalactiae* was inducibly transcribed in the presence of hyaluronan, indicating that streptococcal cells recognize and respond to GAG. Distinct from BacillusUGL, streptococcal UGLs preferred sulfated substrates.

2. Three-dimensional structures of SagUGL wild-type and both substrate-free and -bound D175N mutants were determined by X-ray crystallography. Several residues in the active cleft bind to the substrate, unsaturated chondroitin disaccharide with a sulfate group at the C-6 position of GalNAc residue. The sulfate group is hydrogen-bonded to Ser-365 and Ser-368 and close to Lys-370. The conversion of SagUGL to BacillusUGL-like enzyme via site-directed mutagenesis demonstrated that Ser-365 and Lys-370 are essential for direct binding and for electrostatic interaction, respectively, for recognition of sulfate group by SagUGL. Molecular conversion was also achieved in SagUGL Arg-236 with an affinity for the sulfate group at the C-4 position of the GalNAc residue. These residues binding to sulfate groups are frequently conserved in pathogenic bacterial UGLs, suggesting that the motif "R-//-S*XX*(S)*X*K" is crucial for degradation of sulfated GAGs.

3. The X-ray crystallographic structure of *P. heparinus* HepC was determined at 2.20 Å resolution using single-wavelength anomalous diffraction. This enzyme comprised an N-terminal α/α-barrel domain and C-terminal antiparallel β-sheet domain as its basic scaffold. Superimposition of HepC and heparin tetrasaccharide-bound HepB suggested that an active site of HepC was located in the deep cleft at the interface between its two domains. Three mutants (N240A, Y294F, and H424A) with mutations at the active site had significantly reduced enzyme activity.

4. UGL substrates are classified as 1,3- and 1,4-types based on the glycosidic bonds. The crystal structure of Phep_2830 specific for 1,4-type substrates was determined at 1.35 Å resolution. In comparison with structures of streptococcal and bacillus UGLs, a pocket-like structure and lid

loop at subsite +1 are characteristic of Phep_2830. Site-directed mutations of the pocket and lid loop led to significantly reduced enzyme activity, suggesting that the pocket-like structure and lid loop are involved in the recognition of 1,4-type substrates by UGLs. Key residues comprising the pocket-like structure of Phep_2830 are conserved in UGLs of intestinal bacteroides, suggesting that structural determinants of Phep_2830 for 1,4-specificity promote a better understanding of the complete degradation of heparin by *Bacteroides* species.

ACKNOWLEDGEMENTS

The author deeply appreciates thoughtfulness and critical advice of Dr. Kousaku Murata, Emeritus Professor of Kyoto University and Professor of Faculty of Science and Engineering, Setsunan University. The author also keeps lesson of Dr. Murata “Do stupid study” in his mind.

The author cannot express enough feelings of appreciation to Dr. Wataru Hashimoto, Associate Professor of Graduate School of Agriculture, Kyoto University, for his polite guidance about basis of research activity and writing style, meaningful discussion, and kind encouragement.

The author is grateful to Dr. Bunzo Mikami, Professor of Graduate School of Agriculture, Kyoto University, for his instruction about X-ray crystallography and technical advice.

The author would like to express his gratitude to Dr. Shigeyuki Kawai, Assistant Professor of Graduate School of Agriculture, Kyoto University, for his helpful suggestion.

The author gratefully thanks Dr. Yukie Maruyama, Post-Doctoral Fellow of Graduate School of Agriculture, Kyoto University, for her invaluable discussion, pertinent suggestion, and heartfelt encouragement.

The author is deeply indebted to Dr. Takafumi Itoh, Post-Doctoral Fellow of Graduate School of Agriculture, Kyoto University (currently Associate Professor of Faculty of Biotechnology, Fukui Prefectural University), for his kindly and strictly teaching.

The author would like to thank Dr. Akihito Ochiai, Post-Doctoral Fellow of Graduate School of Agriculture, Kyoto University (currently Assistant Professor of Graduate School of Science and Technology, Niigata University), for giving much advice.

The author would like to show his appreciation to Dr. Kohei Ogura, PhD Student of Graduate School of Agriculture, Kyoto University (currently Post-Doctoral Fellow of Beth Israel Deaconess Medical Center, Harvard Medical School), for his technical support.

The author would like to express his cordial gratitude to Dr. Kazuto Ohashi, PhD Student of Graduate School of Agriculture, Kyoto University (currently Post-Doctoral Fellow of Department of Integrative Structural and Computational Biology, The Scripps Research Institute), for his technical help.

The author is very thankful to Ms. Chizuru Tokunaga, Mr. Takuya Yokoyama, and Ms. Ai Matsunami, Technical Assistants of Laboratory of Basic and Applied Molecular Biotechnology, Graduate School of Agriculture, Kyoto University, for their support.

The author thanks Mr. Ryuichi Takase, PhD Student of Graduate School of Agriculture, Kyoto

University, for exciting discussion.

The author would like to show appreciation to the members of Laboratory of Basic and Applied Molecular Biotechnology, Graduate School of Agriculture, Kyoto University, for their assistance and considerable encouragement.

The author thanks his family for respecting his will and supporting him.

Finally, this study was supported by a Research Fellowship from Japan Society for the Promotion of Science for Young Scientists.

Yusuke NAKAMICHI

LIST OF PUBLICATIONS

Original papers

1. **Yusuke Nakamichi, Bunzo Mikami, Kousaku Murata, and Wataru Hashimoto.** Crystal structure of a bacterial unsaturated glucuronyl hydrolase with specificity for heparin.
   ***J. Biol. Chem.*, 289(8),** 4787-4797 (2014).

2. **Wataru Hashimoto, Yukie Maruyama, Yusuke Nakamichi, Bunzo Mikami, and Kousaku Murata.** Crystal structure of *Pedobacter heparinus* heparin lyase Hep III with the active site in a deep cleft.
   ***Biochemistry,* 53(4),** 777-786 (2014).

3. **Yusuke Nakamichi, Yukie Maruyama, Bunzo Mikami, Wataru Hashimoto, and Kousaku Murata.** Structural determinants in streptococcal unsaturated glucuronyl hydrolase for recognition of glycosaminoglycan sulfate groups.
   ***J. Biol. Chem.*, 286(8),** 6262-6271 (2011).

4. **Yukie Maruyama, Yusuke Nakamichi, Bunzo Mikami, Wataru Hashimoto, and Kousaku Murata.** Substrate specificity of streptococcal unsaturated glucuronyl hydrolases for sulfated glycosaminoglycan.
   ***J. Biol. Chem.*, 284(27),** 18059-18069 (2009).

Related article

1. **Yusuke Nakamichi, Aya Yoshioka, Shigeyuki Kawai, and Kousaku Murata.** Conferring the ability to utilize inorganic polyphosphate on ATP-specific NAD kinase.
   ***Sci. Rep.*, 3,** Article number 2632 (2013).